

Pasientrapporterte utfallsmål for bekkenleddssmerter

Diskriminerende validitet av PGQ, ODI og SF-8

Hanne Krogstad Jenssen



Masteroppgave i Helsefagvitenskap
Institutt for helse og samfunn, Avdeling for helsefag
Det medisinske fakultet

UNIVERSITETET I OSLO

Desember 2012

Pasientrapporterte utfallsmål for bekkenleddssmerter

Diskriminerende validitet av PGQ, ODI og SF-8

Hanne Krogstad Jenssen

Masteroppgave i Helsefagvitenskap
Institutt for helse og samfunn, Avdeling for helsefag
Det medisinske fakultet

UNIVERSITETET I OSLO

Desember 2012

© Hanne Krogstad Jenssen

År 2012

Tittel

Hanne Krogstad Jenssen

<http://www.duo.uio.no/>

Trykk: Reprosentralen, Universitetet i Oslo

Sammendrag

Bakgrunn: Bekken - og korsryggsmerter under graviditeten og etter fødsel er et sammensatt problem med uklare kriterier. Siden det er holdepunkter for at det er vesentlig å skille bekkenleddssmerter fra ryggplager har det vært behov for et sykdomsspesifikt utfallsmål for bekkenleddssmerter. Pelvic Girdle Questionnaire (PGQ) er det første sykdomsspesifikke utfallsmål utviklet for kvinner med bekkenleddssmerter i forbindelse med svangerskap og fødsel. Så langt er PGQ i liten grad validert for diskriminerende evne.

Hensikt: Hensikten med denne studien er å undersøke diskriminerende validitet av PGQ, Oswestry Disability Index (ODI) og SF-8 Short Form (SF-8).

Materiale og metode: Førti-ni kvinner med bekkenleddssmerter og 49 kvinner uten bekkenleddssmerter fylte ut utfallmålene; PGQ, ODI og SF-8. For å teste hypotesen om at kvinner med bekkenleddssmerter ville skåre høyere sammenlignet med kvinner uten bekkenleddssmerter, ble den ikke-parametriske testen Mann-Whitney test benyttet. For å vurdere diskriminerende validitet ble Receiver operating characteristic (ROC) kurver brukt.

Resultater: Analysene viste at kvinnene med bekkenleddssmerter skåret signifikant mye høyere enn kvinnene uten bekkenleddssmerter. ROC kurvene viste at alle instrumentene skilte svært godt mellom kvinner med bekkenleddssmerter og kvinner uten slike plager.

Konklusjon: Masteroppgaven har vist at PGQ, ODI og SF-8 diskriminerer godt mellom kvinner med bekkenleddssmerter og kvinner uten bekkenleddssmerter.

Abstract

Background: Pregnancy–related back pain and pelvic girdle pain is a common condition and give rise to functional disability. In general the literature does not make distinction between pregnancy- related low back pain and pelvic girdle pain. There is, however growing evidence that pelvic girdle pain comprises a distinct subgroup with unique clinical presentation and need for specific management. The Pelvic Girdle Questionnaire (PGQ) is the first condition specific questionnaire developed for women with pelvic girdle pain in pregnancy and postpartum. So far the PGQ is to a minor degree evaluated for discriminative validity.

Objective: The aim of this study was to examine discriminative validity of the PGQ, the Oswestry Disability Index (ODI) and the 8- item version of Medical Outcomes study 36-Item Short-Form Health Survey (SF-8).

Methods: Forty-nine women with pelvic girdle pain and 49 women without pelvic girdle pain completed the outcome measures: PGQ, ODI and SF-8. The non-parametric Mann-Whitney test was used to test the hypothesis that woman with pelvic girdle pain would score higher than woman without pelvic girdle pain on all three instruments. Discriminative validity was assessed with the area under the receiving operating characteristics (ROC) curves.

Results: Women with pelvic girdle pain scored significant higher on the three outcome measures compared to the woman without pelvic girdle pain. The ROC curves showed a high area under the curve for all three instruments.

Conclusion: PGQ, ODI and SF-8 showed all very good discriminative validity between woman with pelvic girdle pain and women without pelvic girdle pain.

Forord

Masterstudiet i helsefag ved UIO har vært en spennende og lærerik tid. Jeg har truffet mange inspirerende og dyktige mennesker.

Jeg har hatt mye støtte og god veiledning til å gjennomføre denne masteroppgaven, og det er flere som fortjener en stor takk.

Først og fremst vil jeg takke min hovedveileder Britt Stuge, uten deg ville jeg aldri ha klart å fullføre denne masteroppgaven. Du har veiledet og støttet meg gjennom hele prosessen. Tusen takk for at du alltid har vært tilgjengelig og hatt god tid til veiledning.

Tusen takk til biveileder Margreth Grotle som har veiledet gjennom hele prosessen. Jeg er veldig takknemlig for ditt fantastiske bidrag i slutfasen av oppgaven.

Tusen takk til mine kollegaer ved Avdeling for fysikalsk medisin og rehabilitering OUS som har støttet og oppmuntret med gjennom hele studietiden. Ekstra takk til min leder Mona Mortensen som har støttet meg fult ut i hele prosessen og ikke minst gitt meg stor fleksibilitet.

En stor takk til Forskningsmiljøet på Ortopedisk senter som har gitt meg enorm støtte og motivasjon til å kunne klare å gjennomføre denne oppgaven. Tusen takk for at jeg har fått tilgang til kontorplass under hele studietiden.

Tusen takk til venner og familie som har støttet og trodd på meg. Nå gleder jeg meg til å være sosial igjen.

Caroline, Einar og Matz uten deres tålmodighet hadde jeg aldri klart å komme i mål med oppgaven. Gleder meg til å være masse sammen med dere igjen. Sist men ikke minst tusen takk Thomas for at du har støttet meg gjennom hele denne prosessen og trått til ekstra på hjemmefronten den siste tiden.

Hanne

Innholdsfortegnelse

1	Innledning.....	1
1.1	Bakgrunn for oppgaven.....	1
1.2	Formålet med oppgaven.....	2
1.3	Oppbygning av oppgaven.....	3
1.4	Begrepsavklaringer.....	3
2	Teoretisk bakgrunn.....	5
2.1	Bekkenleddssmerter	5
2.1.1	Terminologi.....	5
2.1.2	Diagnostisering.....	5
2.1.3	Forekomst.....	7
2.1.4	Etiologi og patogenese	8
2.1.5	Prognose	9
2.2	ICF, Verdens Helseorganisasjons klassifisering av funksjon, funksjonshemming og helse.....	9
2.2.1	Bekkenleddssmerter I lys av ICF	9
2.2.2	ICF som begrepsapparat.....	11
2.3	Pasientrapporterte utfallsmål.....	14
2.3.1	Pelvic Girdle Questionnaire (PGQ).....	15
2.3.2	Oswestry Disability Index Questionnaire (ODI).....	16
2.3.3	SF- 8 Health Survey (SF- 8).....	17
2.3.4	Klassifisering av pasientrapporterte utfallsmål med ICF	18
2.4	Metode for å vurdere kvalitet på pasientrapporterte utfallsmål	19
2.4.1	Reliabilitet	21
2.4.2	Validitet.....	22
2.4.3	Responsiveness – diskriminerende longitudinell validitet	24
2.4.4	Fortolkning	25
2.5	Hoved- og delproblemstillinger	25
3	Metode.....	27
3.1	Studiedesign	27
3.2	Materiale.....	27
3.2.1	Bakgrunn for valg av deltagere	27

3.2.2	Inklusjonskriterier for kvinner med bekkenleddssmerter.....	27
3.2.3	Inklusjonskriterier for kvinner uten bekkenleddssmerter.....	29
3.2.4	Eksklusjonskriterier for begge gruppene.....	29
3.3	Datainnsamling.....	30
3.3.1	Inkludering av kvinner med bekkenleddssmerter	30
3.3.2	Inkludering av kvinner uten bekkenleddssmerter	31
3.4	Pasientrapporterte utfallsmål.....	31
3.4.1	Pelvic Girdle Questionnaire (PGQ).....	31
3.4.2	Oswestry Disability Index, versjon 2.0 (ODI 2.0)	32
3.4.3	SF- 8 Health Survey (SF- 8).....	32
3.5	Databehandling.....	33
3.6	Etisk godkjenning.....	34
4	Resultater.....	35
4.1	Utvalget	35
4.2	Datakvalitet	36
4.3	Fordeling av del- og totalskårer i utfallsmålene.....	39
4.4	Sammenligning av skårer hos kvinner med og uten bekkenleddssmerter	41
4.5	Diskriminerende evne	45
5	Diskusjon.....	47
5.1	Resultater i lys av andre relevante studier.....	47
5.2	Utfallsmål for bekkenleddssmerter	49
5.2.1	PGQ.....	49
5.2.2	ODI.....	51
5.2.3	SF- 8	52
5.3	Resultater i lys av metodiske aspekter	53
6	Konklusjon	59
	Litteraturliste	61
	Vedlegg	69

Figur 1. ICFs modell av funksjon og funksjonshemming (WHO 2001).....	12
Tabell 1: Klassifisering av elementene i PGQ, ODI og SF-8 med ICF	19
Figur 2. COSMIN gruppens anbefalinger ved testing av helse-relaterte pasientrapporterte utfallsmål	20
Tabell 3: Skåringer på hvert enkelt utsagn i de tre instrumentene – PGQ,ODI, og SF-8-	37
Tabell 4: Skåringer på hvert enkelt utsagn i de tre instrumentene -PGQ-ODI og SF-8-	38
Figur 3: Histogrammer med normalfordelingskurve av skåringene i PGQ Aktivitet, PGQ Symptom og ODI	40
Figur 4: Histogrammer med normalfordelingskurve av skåringene SF-8 Fysisk komponent og SF-8 Mental komponent.....	41
Figur 5: Median og interkvartil range for PGQ delskalaene og total skåren	43
Figur 6: Median og interkvartil range for ODI	43
Figur 7: Median og interkvartil range for SF-8 Fysisk komponent og SF-8 Mental komponent	44
Tabell 5: Median skårer og Mann-Whitney test med interkvartil range (IQR) for PGQ Aktivitet skår, PGQ Symptom skår, PGQ Total skår, ODI, SF-8 Fysisk komponent og SF-8 Mental komponent.....	44
Figur 8: Sensitivitet og spesifisitet av Pelvic Girdle Questionnaire (PGQ) Aktivitet, Symptom og Total skår og Oswestry Disability Index (ODI).	45
Figur 9: Sensitivitet og spesifisitet av Short Form – 8 (SF-8) Fysisk komponent og Mental komponent skår.	46
Tabell 6: Areale under ROC kurven	46

Liste over forkortelser

ASLR Active Straight Leg Raise

PGQ Pelvic Girdle Questionnaire

P4 Posterior pelvic pain provocation

ODI Oswestry Disability Index

SF-8 Short form 8 Health Survey

1 Innledning

1.1 Bakgrunn for oppgaven

I min kliniske praksis som spesialfysioterapeut på Oslo universitets sykehus (OUS) har jeg gjennom årene spesialisert meg innen rygg- og bekkenplager. Høsten 2007 ble det åpnet en tverrfaglig poliklinikk for utredning av pasienter med nakke-, rygg- og bekkenplager. Siden åpningen av Nakke- og rygg poliklinikken har jeg hatt ansvaret for fysioterapi for henviste kvinner med bekkenplager, som jeg heretter velger å omtale som bekkenleddssmerter (begrunnes under terminologi senere i oppgaven). I den forbindelse så jeg viktigheten av å ha gode instrumenter for kartlegging og evaluering av denne pasientgruppen. Da det ikke fantes et sykdomsspesifikt funksjonsskjema som var utviklet spesielt for pasienter med bekkenleddssmerter rådet det tvil om hvilket skjema som egnet seg best for denne gruppen.

I tidligere studier av pasienter med bekkenleddssmerter er det benyttet ulike funksjonsskjemaer. Felles for disse skjemaene er at skjemaene er utarbeidet for pasienter med ryggplager eller generelle muskel- og skjelett plager. Et av de mest vanlige er Disability Rating Index (Salèn et al. 1994)) som har blitt brukt i en rekke studier av gravide kvinner med bekkenleddssmerter (Nilsson-Wikmar et al. 2005; Noren et al. 1997; Olsson and Nilsson-Wikmar 2004; Robinson et al. 2010a; Wedenberg et al. 2000). Oswestry Disability Index (Fairbank et al. 1980) som finnes i fire ulike versjoner har også vært benyttet i flere studier av kvinner med bekkenleddssmerter (Gutke et al. 2006; Stuge et al. 2004a)). Videre har Roland & Morris Disability Questionnaire (Roland and Fairbank 2000) og Quebeck Back Pain Disability Scale (Kopce et al. 1995) også blitt brukt for å studere kvinner med bekkenleddssmerter (Bastiaenen et al. 2006; Mens et al. 2002; Padua et al. 2002).

I følge de Europeiske retningslinjene for bekkenleddssmerter er det grunnlag for å skille mellom ryggplager og bekkenleddssmerter. Retningslinjene påpeker at det trengs et evalueringsverktøy som er reliabelt, valid og sensitivt for endringer for pasienter med bekkenleddssmerter (Vleeming et al. 2008). Dette styrket grunnlaget og behovet for å utvikle et sykdomsspesifikt funksjonsskjema for pasienter med bekkenleddssmerter. I et fagutviklingsprosjekt støttet av Fond til etter- og videreutdanning utviklet jeg i samarbeid med fysioterapeutene og forskerne Britt Stuge og Margreth Grotle i perioden 2008-2010 et

sykdomsspesifikt funksjonsskjema som vi har navngitt som The Pelvic Girdle Questionnaire (PGQ).

Arbeidet med å utvikle PGQ resulterte i to nylig publiserte artikler i tidsskriftet Physical Therapy; "The pelvic girdle questionnaire: a condition-specific instrument for assessing activity limitations and symptoms in people with pelvic girdle pain" (Stuge et al. 2011) og "Reliability and construct validity of self-report questionnaires for patients with pelvic girdle pain" (Grotle et al. 2012). PGQ er det første sykdomsspesifikke funksjonsskjema for pasienter med bekkenleddssmerter. PGQ viste seg i disse studiene å ha gode metodiske egenskaper for pasienter med bekkenleddssmerter, både under graviditeten og etter fødsel. Skjemaet er enkelt å administrere og egner seg for bruk i klinisk praksis så vel som i forskning. Det tar ca tre minutter å fylle ut skjemaet og det er enkelt å skåre. Skjemaet har vist god intern konsistens, test-retest reliabilitet og begrepsvaliditet for kvinner med bekkenleddssmerter. Videre har det vist god diskriminerende validitet for forskjellig alvorlighetsgrad av bekkenleddssmerter. Diskriminerende validitet mellom friske og kvinner med bekkenleddssmerter ble imidlertid ikke undersøkt under utviklingen av skjemaet. Dette fordi kun kvinner med bekkenleddssmerter var inkludert i den studien. Vi så derfor et behov for å gjøre en studie for å undersøke diskriminerende validitet.

Anledningen til å gjennomføre en slik studie bød seg i forbindelse med oppstart av bekkenprosjektet "Bekkenleddsplager og bekkenbunn - noen sammenheng?", som var ledet av min hovedveileder Britt Stuge. Til denne bekkenstudien skulle det rekrutteres 50 kvinner med bekkenleddssmerter etter graviditet og 50 friske kontroller. Som en del studie i dette prosjektet fikk jeg da anledning til å undersøke diskriminerende validitet av både det nye PGQ og andre egenrapporterte spørreskjemaer hos kvinner med bekkenleddssmerter og kvinner uten bekkenleddssmerter.

1.2 Formålet med oppgaven

Hensikten med denne masteroppgaven er å sammenligne hvordan kvinner med og uten bekkenleddssmerter skårer på henholdsvis PGQ, ODI og SF-8. Primært vil instrumentenes diskriminerende evne bli undersøkt, men også data kvalitet som skåringenes fordeling i forhold til skalaen, samt mulige gulv- og takeffekter.

1.3 Oppbygning av oppgaven

Masteroppgaven er en metodestudie der PGQ, ODI og SF-8 blir testet for diskriminerende validitet mellom kvinner med og uten bekkenleddssmerter.

I teorikapittelet vil jeg redegjøre for bekkenleddssmerter, både med hensyn til terminologi, diagnose og forekomst. Jeg vil også kort redegjøre for dagens kunnskaper angående etiologi og patogenese, men ikke komme inn på behandling. Deretter følger et kapittel om kartlegging og evaluering av bekkenleddssmerter, og jeg presenterer dette i lys av referanserammen ICF, Verdens Helseorganisasjons klassifikasjon av funksjon, funksjonshemming og helse.

Pasientrapporterte utfallsmål som er benyttet i denne masteroppgaven presenteres før jeg gir en oversikt over metodiske kriterier for å vurdere kvalitet på de utfallsmål som jeg vurderer. I midtre delen av oppgaven presenterer jeg metode og resultater. Deretter diskuteres metode og resultater før jeg konkluderer.

1.4 Begrepsavklaringer

Intern konsistens (internal consistency) omhandler i hvor stor grad spørsmålene i et utfallsmål er relatert til hverandre.

Målefeil (measurement error) har med den absolutte målefeilen, det vil si hvor liten avstand det er mellom de repeterte målingene.

Reliabilitet (reliability) handler om i hvilken grad en målemetode er fri for målefeil.

Innholdsvaliditet (content validity) omhandler i hvilken grad det man ønsker å måle blir fanget opp av utsagnene i et utfallsmål.

Kriterium validitet (criterion validity) referer til i hvilken grad et utfallsmål er relatert til en gullstandard.

Begrepsvaliditet (construct validity). I forskningen kalles et fenomen som man forsøker å måle med tall eller kategorier for et "construct". Det finnes ingen god oversettelse av dette

begrepet, men noen anvender begrepsvaliditet. Begrepsvaliditet omhandler i hvilken grad resultatene av måleinstrumentet er konsekvent med hypotesene med hensyn til interne forhold, relasjoner med en rekke andre instrumenter eller forskjellige relevante grupper.

Fortolkning (interpretability) omhandler i hvilken grad man kan tillegge kvalitativ mening til kvalitativ skår.

Gulv – takeffekt (floor-ceeling effect) kan oppstå når en høy andel av den totale populasjonen har skåringer på den laveste eller øverste delen av skalaen.

Responsivness blir definert som et utfallsmåls evne til å fange opp endring over tid når en endring har skjedd.

2 Teoretisk bakgrunn

2.1 Bekkenleddssmerter

2.1.1 Terminologi

Bekken- og korsryggsmerter under graviditeten er et sammensatt problem med uklare kriterier og blir ofte omtalt som bekkenløsning (Endresen 1995). Bekkenløsning er imidlertid en naturlig fysiologisk prosess under svangerskap som ikke skal gi smerter og funksjonshemming (Stuge 2010). Hvis kvinnen derimot opplever smerter av funksjonshemmende karakter og melder seg som pasient oppfattes det som en patologisk tilstand (Kogstad and Biornstad 1990). Denne tilstanden har gjennom tidene hatt ulike betegnelser på bakgrunn av flere mulige årsaksforklaringer og symptomer. Eksempler på betegnelser som internasjonalt er brukt for denne tilstanden er; pelvic girdle relaxation, symptom-giving pelvic relaxation, pelvic joint instability, pubic symphysis diastasis og symphyseal separation (Bjelland et al. 2012). De mange og forskjellige betegnelsene reflekterer at smerter i bekkenet er en tilstand med usikker underliggende patologi og at smertelokalisasjon kan variere. Betegnelsen bekkenløsning kan imidlertid gi et inntrykk av at noe er løst og det kan gi et uheldig bilde for pasienten. I den senere tid har man i Norge ofte brukt begrepet bekkenleddsplager for bekkenrelaterte smerter som ekskluderer plager av gynekologisk og urologisk karakter (Stuge 2010). Imidlertid diskuteres det hvorvidt bekkenleddsplager eller bekkenleddssmerter vil være den beste betegnelsen relatert til smerter i muskelskjelettsystemet rundt bekkenet. I de Europeiske retningslinjene benyttes betegnelsen "Pelvic Girdle Pain" (Vleeming et al. 2008). Ordet 'girdle' ble bevisst tatt med for å tydeliggjøre at det her gjelder smerter fra muskel-skjeleett systemet, og ikke gynekologiske lidelser som dekkes av betegnelsen 'pelvic pain'. I denne oppgaven velger jeg å bruke betegnelsen bekkenleddssmerter, da smerte er den direkte oversettelsen av pain.

2.1.2 Diagnostisering

Det finnes imidlertid ikke en gullstandard for diagnostisering av bekkenleddssmerter. Det er derfor viktig med en grundig anamnese for å kartlegge smertens lokalisering, alvorlighetsgrad

og karakter, provoserende faktorer og funksjonsproblemer. I tillegg fremheves det i de Europeiske retningslinjene at kliniske tester er nødvendige for å kunne stille en diagnose (Vleeming et al. 2008). Smertetegning kan også være et nyttig hjelpemiddel. Et av målene med diagnostisering av bekkenleddssmerter kan være å skille mellom ryggplager og bekkenleddssmerter. Bekkenleddssmerter er som regel lokalisert til sacrum, under spina iliaca posterior superior, i gluteal regionen, dorsalside av lår og/eller symfysen. Andre regioner som ofte er affisert er lysken og halebenet (Stuge 2010; Vleeming et al. 2008). Ryggsmerte derimot er som regel lokalisert til lumbalregionen over sacrum eller defineres som smerter mellom 12. ribbe og gluteal folden (Stuge 2010; Vleeming et al. 2008). I utvikling av PGQ fant vi at det å gå er en av de aktivitetene som er mest provoserende for kvinner med bekkenleddssmerter (Stuge 2010). For øvrig fremheves det å sitte og stå som aktiviteter som er provoserende over tid (Vleeming et al. 2008). Følelsen av at benet svikter er et annet diagnostisk tegn for bekkenleddssmerter (Stuge et al. 2004b; Stureson et al. 1997).

For best mulig å diagnostisere eller klassifisere bekkenleddssmerter er det viktig med reliable og valide kliniske tester i tillegg til anamnesen. Det finnes imidlertid i dag ingen klinisk test som er tilstrekkelig for å diagnostisere bekkenleddssmerter. Europeiske retningslinjer for bekkenleddssmerter anbefaler at man bruker flere kliniske tester når man vurderer denne pasientgruppen (Vleeming et al. 2008). Tester som anbefales er Posterior pelvic pain provocation test, Kompresjonstest, Distraksjonstest, Patrick Fabere test, palpasjon av de lange dorsale iliosacralleddsligamentene, symfyse tester og Active Straight Leg Raise (ASLR) test. En av testene som er anbefalt for provokasjon av iliosacralleddene er Posterior pelvic pain provocation (P4) test. Dette er en smerteprovokasjonstest som har vist seg som en god diagnostisk test og benyttes både i forskning og klinisk praksis (Ostgaard et al. 1994). Denne testen gir informasjon om hvor smerten er lokalisert, men ikke hvilke anatomiske strukturer smerten kommer fra. Kompresjonstest skal angivelig gi strekk på bakre iliosakrale ligamentet og kompresjon av fremre del av iliosacralleddet, mens Distraksjonstest er en test som skal gi strekk av de fremre iliosakrale ligamenter og gi kompresjon dorsalt i leddet (Laslett et al. 2005). Patrick-Fabre test er også en test som kan provosere frem smerter i iliosacralleddet (Laslett et al. 2005a; Vleeming et al. 2008). Videre er test av de lange dorsale iliosacralleddsligamentene en palpasjons test for å provosere frem smerter i området rundt iliosacralleddene (Vleeming et al. 2002). For symfyse plager anbefales palpasjon av symfysen og modifisert Trendelenburg test (Albert et al. 2002). ASLR derimot, er en funksjonstest for å

vurdere evnen til å overføre krefter mellom trunkus og bena og er mye benyttet i studier for å undersøke motorisk kontroll (Beales et al. 2009a; Beales et al. 2009b; Mens et al. 1999).

De aller beste testene for å identifisere bekkenleddssmerter har vist seg å være P4-testen og ASLR testen, som begge har en sensitivitet og spesifisitet på 80 % (Vleeming et al. 2008). Testene skiller godt mellom de med og de uten bekkenleddssmerter og P4-testen har vist seg å være negative hos pasienter med ryggplager (Gutke et al. 2009). De kliniske testene vil bli beskrevet nærmere under avsnittet om datainnsamling.

2.1.3 Forekomst

Bekkenleddssmerter i forbindelse med svangerskap og etter fødsel er et problem for mange kvinner. De fleste studier har vist en forekomst på 45 % av rygg og/eller bekkenleddssmerter i svangerskapet med en variasjon fra 4 % til 90 % (Wu et al. 2004). Europeiske retningslinjer for bekkenleddssmerter konkluderer med at det er sterk evidens for en forekomst på 20 % for bekkenleddssmerter i svangerskapet (Vleeming et al. 2008). I henhold til Wu og medarbeidere (Wu et al. 2004) har 25 % rygg og/eller bekkenleddssmerter etter fødsel, men også her rapporteres stor variasjon (0.3 % til 67 %). I følge Wu's beregninger er det ca 5 % av disse kvinnene som har alvorlig grad av bekkenleddssmerter. Det er ulike årsaker til at det i litteraturen finnes så stor variasjon i forekomst. En av årsakene er at noen av studiene på forekomst er prospektive og andre er retrospektive og at data i stor grad er samlet inn gjennom spørreskjema. Ulike diagnostiske kriterier er også med på å påvirke forekomst tallene. Noen mener at bekkenleddssmerter er en undergruppe av diffuse ryggplager, mens andre hevder at det er vesentlig å skille bekkenleddssmerter både diagnostisk og behandlingsmessig fra ryggplager (Albert et al. 2000; Ostgaard et al. 1994). I en nylig publisert studie av Robinson og medarbeidere (2010a) går det frem at gravide skårer ulikt på funksjon avhengig av om de har bekkenleddssmerter eller ei.

2.1.4 Etiologi og patogenese

Kunnskap om årsakene til bekkenleddssmerter er mangelfull. Hormonelle og biomekaniske faktorer er mulige årsaksforklaringer, men reliabel kunnskap om mekanismene mangler (Kogstad and Biørnstad 1990; Stuge 2010). Svangerskapshormonet relaksin er vist å ha en medvirkende rolle i remodellering av bekkenleddsstrukturene i svangerskapet, men flere studier har ikke funnet positiv sammenheng mellom relaksin og bekkenleddssmerter (Aldabe et al. 2012b). Bekkenleddssmerter kan debutere så tidlig som 4-5 svangerskapsuke, og symptomene forsvinner hos de fleste kvinner kort etter fødsel (Bjelland et al. 2012). Dette kan indikere at bekkenleddssmerter ikke har noen sammenheng med økt vekt og endret tyngdeplassering, men heller kan skyldes en hormonell påvirkning. I en nylig publisert artikkel fant Bjelland og medarbeidere (2011) at tidlig menarke hadde sammenheng med bekkenleddssmerter og det kan tyde på at hormonelle faktorer kan være en mulig årsaksfaktor. Vøllestad og medarbeidere (2012) fant imidlertid ikke dokumentasjon for at relaksin har innvirkning på symptomer og funksjon under graviditet.

Ligamentstrukturene rundt iliosakralleddene og området rundt festet til det sacrospinale ligamentet kan være en annen kilde til bekkenleddssmerter. Torstensson og medarbeidere (2009) konkluderte i en studie at dette er et område som kan være kilde til langvarige lumbosakrale smerter som kan oppstå under graviditeten. Overfladiske iliosacralleddsstrukturer som mulig smertekilde for bekkenleddssmerter bekreftes i en nylig publisert studie (Palsson and Graven-Nielsen 2012).

I en systematisk oversikt over mobilitet av bekkenledd ble det konkludert med at det var større mobilitet i bekkenleddene under svangerskapet og etter fødsel hos kvinner med bekkenleddssmerter enn hos friske kontrollpersoner (Mens et al. 2009). Det ble likevel presisert at klinikere ikke måtte bruke symfyse åpning og bevegelse i bekkenleddene som diagnostiske kriterier da det er stor variasjon i mobilitet. Selv om økt mobilitet i bekkenleddene ikke har vist en klar sammenheng mellom smerter i bekkenet, har derimot asymmetrisk mobilitet vist seg å ha en sammenheng med bekkenleddssmerter (Damen et al. 2001). Det kan tenkes at asymmetrisk mobilitet og økt mobilitet i bekkenleddene kan påvirke motorisk kontroll som også hevdes å være en medvirkende årsak til bekkenleddssmerter (Mens et al. 1999; Stuge 2010). Aldabe og medarbeidere (2012a) konkluderer i en nylig publisert oversiktsartikkel med moderat evidens for at bekkenleddssmerter kan være relatert til motorisk kontroll og mekaniske faktorer.

2.1.5 Prognose

Bekkenleddssmertene forsvinner kort tid etter fødselen hos de aller fleste, men noen kvinner rapporterer vedvarende smerter lenge etter fødsel. Det er imidlertid mangelfull kunnskap om prognose av bekkenleddssmerter etter fødsel blant annet fordi det finnes få longitudinelle studier. Det er også vanskelig å sammenligne studiene da de i hovedsak har benyttet ulike design og utfallsmål. I tillegg er det få studier som har sett på langtidsoppfølging (Albert et al. 2001; Bjelland et al. 2012). En nylig publisert studie viser at nærmere 80 % har blitt bra av sine selvrapporterte bekkenleddssmerter seks måneder etter fødsel, mens ca 3 % rapporterer betydelige smerter (Bjelland et al. 2012). I henhold til de Europeiske retningslinjene er det ca 7 % som opplever vedvarende bekkenleddssmerter etter fødsel (Vleeming et al. 2008). Studier har for øvrig vist at kvinner med smerte i både bekkenet og korsrygg tidlig i svangerskapet, høy smerte skåring og smerter lokalisert i alle tre bekkenleddene er faktorer som gir dårligere prognose (Albert et al. 2001; Bjelland et al. 2012; Gutke et al. 2008). Robinson og medarbeidere (2010b) har videre vist at mange positive provokasjons tester er risikofaktorer for både nedsatt funksjonsevne og smerter 12 uker etter fødsel. Det er også vist at positiv test av aktiv strakt beinløft (ASLR) og troen på bedring kan predikere bedring hos kvinner med bekkenleddssmerter etter fødsel (Vollestad and Stuge 2009).

2.2 ICF, Verdens Helseorganisasjons klassifisering av funksjon, funksjonshemming og helse

2.2.1 Bekkenleddssmerter I lys av ICF

Etiologi og patogenese av bekkenleddssmerter er som nevnt mangelfull (Bjelland 2012 s. 11, Kogstad 1990). Kvinner med bekkenleddssmerter kan derfor presentere symptomer uten en tydelig manifestasjon av underliggende patologi og kan på tross av det ha stor grad av funksjonshemming. Studier har vist at det finnes mange pasienter som kommer til helsevesenet med symptomer uten underliggende patologi (Carson et al. 2000). Wade og Halligan (2004) hevder at pasienter med symptomer som i liten grad kan forklares på bakgrunn av underliggende patologi, ofte får en medisinsk diagnose som impliserer en

underliggende strukturell årsak og reflekterer kulturelle forventninger. Forklaringsmodellen for sykdom blir da basert på den biomedisinske modellen som har vært den rådende forklaringsmodellen gjennom det 20. århundre. Denne modellen er basert på Virchows konklusjon om at all sykdom er et resultat av cellulære uregelmessigheter (Wade and Halligan 2004) og har blitt forbundet med antagelser om at alle symptomer og plager er avledet fra en underliggende uregelmessighet i kroppen (vanligvis funksjon eller struktur i et spesifikt organsystem). I den medisinske modellen er det ikke rom for sosiale, psykologiske eller adferdsmessige dimensjoner av sykdom (Engel 1977) og sykdom blir sett på som;

”funksjonshemming som et individuelt problem, forårsaket av sykdom, skade eller annen helsetilstand som trenger medisinsk behandling i form av individuell behandling utført av helsepersonell” (Sosial og helsedirektoratet 2003). Den sosiale modellen derimot ser funksjonshemming som et sosialt skapt problem og prinsipielt som et spørsmål om individets fullstendige integrering (Haugsbø and Larsen 2004). WHO omtaler den sosiale modellen som;

”en forståelse som er skapt av sosiale forhold og handler grunnleggende om full integrasjon av enkeltpersoner i samfunnet. Da er funksjonshemming ikke en egenskap ved en enkelt person, men snarere en komplisert blanding av tilstander, hvorav mange er skapt av det sosiale miljøet. Derfor krever håndteringen av problemet sosiale tiltak på det samfunnsmessige nivå.” (Sosial og helsedirektoratet 2003)

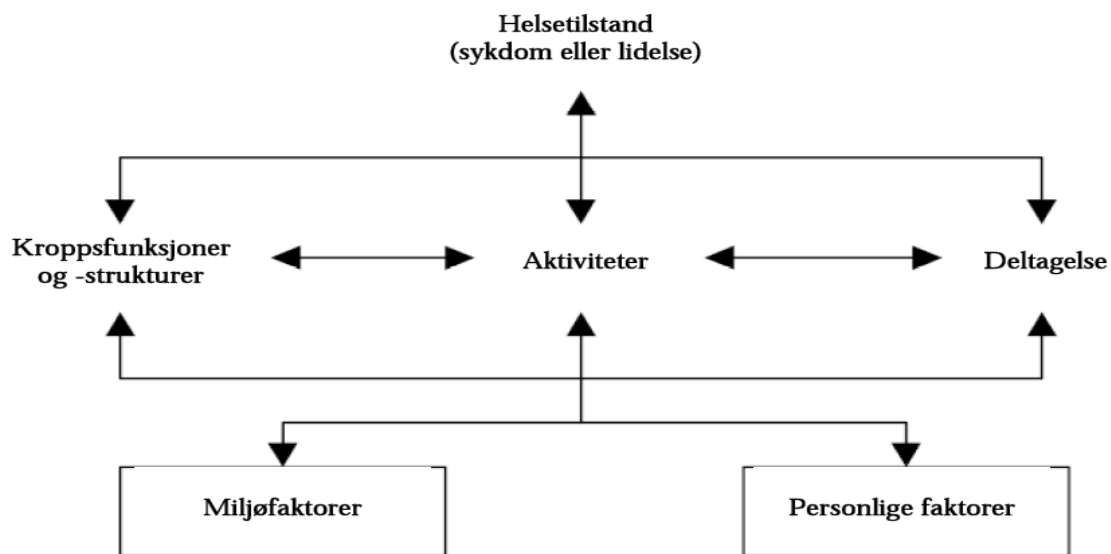
En sammensetting av de to motsatte modellene blir sett på som den biopsykosiale modellen og ble først introdusert av psykiateren Georg Engel (Engel 1977). Han fant den biomedisinske modellen utilstrekkelig for å forstå hvordan sykdom kan oppstå og mente det var viktig å ta med den sosiale omverdenen pasienten lever i for å kunne møte pasientens sykdom. Modellen blir ofte omtalt som relasjonell, da helse i en slik modell blir sett i sammenheng med den fysiske og sosiale omverdenen personen lever i. Den biospsykososiale modellen ser sykdom og helse i et systemperspektiv som er forbundet med et hierarki som går fra det molekylær-biologiske via det psykologiske til det sosiokulturelle nivået. Dermed kan påvirkninger i ett system også utløse forandringer, eventuelt sykdom eller helse i andre systemer (Vaglum 2007). WHO har bidratt til å legitimere denne modellen ved at de i 2001 vedtok Internasjonal klassifikasjon av funksjon, funksjonshemming og helse (ICF) som bygger på en integrering av den biomedisinske - og sosiale modellen. ICF bygger på den tidligere modellen International Classification of Impairments, Disability and Handicaps ICDH fra 1980, som ble kritisert for å være avledet fra den medisinske modellen (Grue 2001). På bakgrunn av erfaringer og utprøvinger ble fokuset ved utviklingen av ICF endret fra å gjelde konsekvenser

av sykdom til å gjelde bestanddeler av helse (Sosial og helsedirektoratet 2003). ICF beskriver faktorer av helsemessige betydning i menneskers livssituasjon, men forholder seg likevel nøytral i årsakssammenhenger og er ikke beregnet for å klassifisere risikofaktorer for helse, eller årsaker til helseskade. Målsettingen for ICF er å gi et grunnlag for vitenskapelige undersøkelser av helse og helserelaterte tilstander og være et felles språk for beskrivelse av helse og helserelaterte forhold. Med dette ønsker man å bedre informasjonsutveksling mellom forskjellige aktører (helsearbeidere, forskere, politikere og offentligheten, inklusive personer med funksjonshemming), fremme utveksling av informasjon på tvers av landegrenser, mellom ulike helsefag og over tid, samt utgjøre et systematisk kodeverk for helseinformasjonssystemer.

2.2.2 ICF som begrepsapparat

ICF er en flerdimensjonal og rasjonell modell som bygger på et syn om at funksjon ikke er entydig, men kan forstås i lys av flere dimensjoner. ICF er delt inn i to hovedområder; funksjon og funksjonshemming, og kontekstuelle faktorer (Sosial og helsedirektoratet 2003). Funksjon og funksjonshemming er delt inn i emneområdene; kroppsfunksjoner og kroppsstrukturer og aktivitet og deltagelse. Kontekstuelle faktorer er delt inn i emneområdene miljø- og personlige faktorer (Fig 1). ICF kan på den ene siden brukes for å påvise funksjonsavvik, aktivitetsavvik eller deltagelsesinnskrenkning som kan omtales som funksjonshemming. På den andre siden kan problemfri helse og helserelaterte forhold beskrives som funksjon. I ICF blir funksjon og funksjonshemming fortolket ved hjelp av begrepskonstruksjoner som operasjonaliseres ved hjelp av såkalte modifikatorer.

Figur 1. ICFs modell av funksjon og funksjonshemming (WHO 2001)



Kroppsfunksjoner og kroppsstrukturer

Kroppen i ICFs begrepsapparat er hele menneskeorganismen inklusive hjernen og dens funksjoner. Det skilles ikke mellom kropp og sinn, derfor er mentale (eller psykiske) funksjoner regnet med blant kroppsfunksjonene (Sosial og helsedirektoratet 2003).

Kroppsfunksjoner er definert som organsystemets fysiologiske funksjoner, mens kroppsstrukturer er de anatomiske delene av kroppen. Modifikatorer for denne delklassifikasjonen er avvik. Avvik i kroppsfunksjoner kan omfatte endret bevegelsesmønster på grunn av smerter i bekkenet og følelse av at benet svikter. Svangerskapshormonet relaksin er vist å ha en medvirkende rolle i remodellering av bekkenleddsstrukturene og dette vil da kunne klassifiseres under kroppsstrukturer. Kroppsfunksjoner påvirkes som andre funksjoner av det yremiljøet som for eksempel hvordan støy kan påvirke hørselen og synet av lysforhold.

Aktivitet og deltagelse

Aktivitet er menneskets utførelse av oppgaver og handlinger. Deltagelse er definert som engasjement i en livssituasjon (Sosial og helsedirektoratet 2003). Aktivitet og deltagelse kan utdypes ved to begrepskonstruksjoner, modifikatorene for utførelse og kapasitet. Utførelse beskriver hva en person gjør i sitt nåværende miljø og innebefatter fysiske, sosiale og

samfunnsmessige forhold. Dette kan også forstås som engasjement i en livssituasjon eller menneskets opplevde erfaringer i deres virkelige livssammenheng. Begrensning eller innskrenkning bedømmes ut fra en allment akseptert standard for befolkningen som ikke har det aktuelle helseproblemet. Kvinner med bekkenleddssmerter klager ofte over smerter når de går står og ligger (Moen et al. 1990) som igjen medfører nedsatt funksjon i disse aktivitetene. Dette vil da bli klassifisert som nedsatt utførelse av aktivitet og pasientrapporterte utfallsmål er en måte å måle dette på. Kapasitet beskriver en persons evne til å utføre en oppgave eller en handling innenfor et bestemt domene til et gitt tidspunkt og bygger på en forståelse om et standardisert miljø. Dette bør da måles ved objektive tester hvor man kan standardisere miljøet. Kapasitet kan kun kodes for aktiviteter, da det er vanskelig å beskrive et standardisert miljø for deltagelse.

Kontekstuelle faktorer

Kontekstuelle faktorer utgjør bakgrunnen for en persons liv og livsutfoldelse. De er delt inn i miljøfaktorer og personlige faktorer (Sosial og helsedirektoratet 2003). Miljøfaktorene beskriver de fysiske, sosiale og holdningsmessige omgivelser hvor en person eksisterer og utfolder sitt liv. Disse faktorene er utenforliggende i forhold til individet og kan ha en positiv eller negativ innvirkning på individets funksjon. De individuelle faktorene omfatter en persons nærmiljø, inklusive hjemmet, arbeidsmiljøet og skole. Materielle særtrekk, samt direkte kontakt med familie, kjente og fremmede inngår her. De samfunnsmessige faktorene omfatter formelle og uformelle sosiale strukturer, tjenester, normer og systemer som har innvirkning på enkeltmennesker i samfunn eller en kultur. Dette inkluderer organisasjoner, samfunnsaktiviteter, offentlige myndigheter, kommunikasjons- og transporttjenester og uformelle sosiale nettverk, lover og forskrifter, formelle og uformelle regler, holdning og ideologier.

Emneområdet personlige faktorer er den spesielle bakgrunnen for en persons eksistens og livsutfoldelse, og omfatter særtrekk ved personen som ikke hører til noen helsetilstand eller helseforhold (Sosial og helsedirektoratet 2003). Disse faktorene kan omfatte kjønn, rase, alder, andre helsetilstander, fysisk form, livsstil, vaner, oppdragelse, alment adferdsmønster, karakteregenskaper og psykiske resurser. Personlige faktorer klassifiseres ikke i ICF, som begrunnes med at det er store kulturelle og samfunnsmessige variasjoner i oppfattelse av

disse. Personlige faktorer er likevel tatt med i figur 2 for å illustrerer hvordan de kan innvirke på ulike tiltak (Sosial og helsedirektoratet 2003).

2.3 Pasientrapporterte utfallsmål

Aktivitet og deltagelse kan måles ved pasientrapporterte utfallsmål. Pasientrapporterte utfallsmål er i dag sterkt anbefalt når man evaluerer nytte eller effekt av diverse behandlingstiltak i kliniske forsøk, rutinemessig kvalitetssikring og registrerings system (Bombardier 2000; Morkink et al. 2010; Storheim et al. 2012). Gode utfallsmål for å kartlegge og evaluere funksjon i klinisk praksis er også viktig. I kliniske studier er gode, pålitelige og valide utfallsmål viktig når man vil måle effekt av behandling eller det kliniske forløpet hos pasienter over tid. Det har i mange år vært tatt som en selvfølge at objektive instrumenter som f.eks, måler ledd bevegelighet og muskelkraft er pålitelige og gyldige, mens påliteligheten har vært omdiskutert når det gjelder pasientrapporterte utfallsmål, slik som spørreskjemaer. Bruk av pasientrapporterte utfallsmål har imidlertid økt voldsomt i omfang de siste ti-årene, og det finnes et stort utvalg av instrumenter som er utviklet for pasienter med rygg smerter (Grotle et al. 2004b). Internasjonale ekspert panel for ryggplager, har foreslått at et "core set" av utfallsmål bør inneholde følgende fem domener; rygg spesifikke funksjonsskjema, generiske helsestatus skjemaer, smerte, arbeidsstatus, og pasient tilfredshet (Bombardier 2000; Deyo et al. 1998). I såkalte 'minimal core set' er det et krav at det skal være med et skjema for selvrapportert funksjon (Bombardier 2000; Deyo et al. 1998; Storheim et al. 2012).

Pasientrapporterte utfallsmål kan enten være generiske eller sykdomsspesifikke. Generiske instrumenter er utviklet for å måle helse- og/eller funksjonsstatus uavhengig av sykdom, mens sykdomsspesifikke instrumenter er laget for å være sensitive for en spesifikk sykdom (Grotle et al. 2004a). Ved valg av utfallsmål må man undersøke i hvilken grad skjemaet er kvalitetssikret i forhold til reliabilitet, validitet og responsiveness (Kjeken et al. 2007). Det er også viktig at skjemaet er anvendelig til de forhold det skal brukes til, enten det er klinisk praksis eller forskning. Jeg vil i det følgende kort beskrive de 3 utfallsmålene som er benyttet i denne oppgaven; PGQ, ODI og SF-8.

2.3.1 Pelvic Girdle Questionnaire (PGQ)

PGQ er et nytt sykdomsspesifikt funksjonsskjema utviklet for kvinner med bekkenleddssmerter (Grotle et al. 2012; Stuge et al. 2011). Skjemaet består av 20 funksjonsspørsmål og fem symptomspørsmål, som kan skåres i en funksjon- og symptom delskala eller i en totalskåre. Før utvikling av PGQ ble det gjort et grundig systematisk litteratursøk og gjennomgang av litteratur for å vurdere hvorvidt det var behov for et nytt instrument. Siden det ikke fantes et sykdomsspesifikt utfallsmål for kvinner med bekkenleddssmerter ble det besluttet å utvikle et slikt instrument. I utviklingen benyttet man seg av potensielle relevante utsagn (item) i de allerede eksisterende og benyttede utfallsmålene, for eksempel ODI, Disability Rating Index (DRI) og Roland Morris Disability Questionnaire (RMDQ). Det ble arrangert fokusgruppe intervjuer med erfarende fysioterapeuter innen feltet og kvinner med bekkenleddssmerter både under graviditet og etter fødsel. To runder med datainnsamling og Rasch analyse ble gjort for å sikre god kvalitet av PGQ. Det første PGQ utkastet inneholdt 40 funksjonsspørsmål og 7 symptom spørsmål. Ved hjelp av Rasch analyse ble antall utsagn redusert ned til 20 funksjonsspørsmål og 5 symptomspørsmål. Rasch analysene viste også at det var vanskelig å skille mellom nivå 4 og 5 på en 5 delt skala. Derfor ble skalaen reduserte ned til en 4 del skåring (Stuge et al. 2011). I den andre datainnsamlingen ble nåværende PGQ versjon med 20 funksjonsspørsmål og 5 symptomspørsmål testet ut og de metodiske egenskapene ble sammenlignet med andre funksjonsinstrumenter slik som ODI og DRI, men også andre psykologiske instrumenter (to instrumenter som måler katastrofetenkning ved smerte, the Pain Catastrophising Scale, og Fear-Avoidance Beliefs Questionnaire). I tillegg ble livskvalitetsinstrumentet SF-8 benyttet. DRI og ODI ble valgt fordi de er de mest brukte funksjonsskjemaene hos kvinner med bekkenleddssmerter, og SF-8 versjonen ble benyttet fordi det var den korteste versjonen av SF-36, som er det mest brukte helserelaterte livskvalitetsinstrumentet.

PGQ har vist gode resultater med hensyn til metodiske egenskaper (Stuge et al. 2011; Grotle et al. 2012). Intern konsistens målt ved Cronbach's alpha var 0.86 på aktivitetsspørsmålene og 0.68 på symptomspørsmålene. Det var verken tak- eller gulveffekt i PGQ skåringene (Grotle et al. 2012). Test-retest reliabiliteten ble testet ut på 42 pasienter med bekkenleddssmerter (21 gravide og 21 ikke gravide). Intraclass correlation coefficient (ICC) var svært høy (>0.90) for begge delskalaene i PGQ; 0.93 for de 20 funksjonsspørsmålene og 0.91 for de fem symptomspørsmålene. Begrepsvaliditet (construct validitet) ble bekreftet av både Rasch

analysen, som viste at PGQ skalaene var unidimensjonale, og av klassisk undersøkelse av hypoteser om forventet korrelasjon av PGQ med andre utfallsmål. Funksjonsskalaen til PGQ, PGQ- aktivitet, viste som forventet en høy korrelasjon med DRI og ODI, mens symptomskalaen, PGQ-symptom, viste en moderat korrelasjon med DRI og ODI. Begge PGQ-delskalaene viste seg å ha bedre diskriminerende evne enn de andre instrumentene med hensyn til skåringene hos gravide og ikke-gravide kvinner samt skåringer hos kvinner med og uten smerte i flere områder i bekkenregionen. Foreløpig er ikke responsiveness av PGQ blitt undersøkt (Grotle et al. 2012).

2.3.2 Oswestry Disability Index Questionnaire (ODI)

Oswestry Disability Index Questionnaire (ODI) er et sykdomsspesifikt spørreskjema som er mye brukt og anbefalt for å evaluere funksjonsgrad hos pasienter med korsryggsmarter (Bombardier 2000; Fairbank and Pynsent 2000; Roland and Fairbank 2000). Det har også blitt brukt som utfallsmål i studier på kvinner med bekkenleddssmerter (Gutke et al. 2006; Stuge et al. 2004a). Utviklingen av ODI ble først igangsatt av Johan O'Brien i 1976 i en spesialist klinikk for pasienter med kroniske ryggsmarter (Fairbank et al. 1980). Utvelgelsen av utsagn/spørsmål var basert på intervju av pasienter med kroniske ryggsmarter angående restriksjoner i forskjellige dagliglivsaktiviteter (Fairbank et al. 1980). Spørreskjemaet ble første gang publisert i 1980 og har blitt modifisert en rekke ganger etter det (Beurskens et al. 1996; Fairbank et al. 1980; Resnik and Dobrykowski 2005). Det er minst fire versjoner av ODI (Grotle et al. 2003). De modifiserte versjonene har blitt vurdert av de opprinnelige forfatterne, og de anbefaler at man bruker versjon 2 av ODI (Baker et al. 1990), som også har blitt oversatt og testet for norske pasienter (Grotle et al. 2003). Denne versjonen inneholder 10 spørsmål som er relatert til daglige aktiviteter og/eller smerte som skåres på en 6 punkts skala. Total skåren summeres og prosentueres til en 0-100 score der 0 representerer ingen nedsatt funksjon og 100 representerer maks nedsatt funksjonsevne.

ODI har vist seg å være et reliabelt og valid instrument for å vurdere funksjonsproblemer hos pasienter med korsryggsmarter i mange studier (Fairbank and Pynsent 2000; Fairbank et al. 1980; Roland and Fairbank 2000). ODI 2.0 versjonen har også vist seg å være spesielt godt egnet for å måle funksjonsproblemer hos norske ryggpasienter når det har blitt sammenlignet med andre sykdomsspesifikke spørreskjema (Grotle et al. 2004a). Flere studier har vist at ODI

har god responsiveness når det er brukt hos pasienter med lumbale ryggsmarter (Beurskens et al. 1996; Fairbank and Pynsent 2000; Resnik and Dobrykowski 2005).

2.3.3 SF- 8 Health Survey (SF- 8)

Spørreskjemaer for generell helsestatus er utviklet for å undersøke bredt i forhold til helse, utførhet og livskvalitet. Selv om generelle helsestatusinstrumenter ofte har vist seg å være mindre responderende for endring enn sykdomsspesifikke instrumenter (Grotle et al. 2004b), så er de viktige for å kunne sammenligne helsestatus i ulike sykdomsgrupper eller i den friske delen av befolkningen (Lurie 2000). Av fem instrumenter for generell helsestatus som ble vurdert av Bombardier og medarbeidere, har SF-36 blitt pekt ut som det beste med tanke på reliabilitet, validitet, og responsiveness når det er testet ut i store populasjoner med korsryggsmarter (Bombardier 2000; Deyo et al. 1998; Lurie 2000). SF-36 er oversatt til 120 språk og administrert til millioner av mennesker verden over (Burholt and Nash 2011). SF-36 inneholder 36 spørsmål/utsagn som presenteres i åtte subskalaer; generell helse, fysisk funksjon, fysisk rollefunksjon, kroppssmerte, overskudd eller vitalitet, sosial funksjon, mental helse og emosjonell rollefunksjon. Man kan summere tre av subskalaene (fysisk funksjon, fysisk rollefunksjon, kroppssmerte) til en fysisk komponent skåre; Physical Component Summary (PCP)(Ware et al. 2001)samt tre av de mentale subskalaene (mental helse, emosjonell rollefunksjon og sosial funksjon) til en mental komponent skåre; Mental Component Summary (MCS).

SF-8 er en kortversjon av SF-36 og inneholder ett spørsmål for hver av de åtte subskalaene, som kan summeres i de samme fysiske og mentale komponent skårene som i SF-36. Målet med å lage kortversjonen SF-8 var at man skulle tilby et måleinstrument som kunne administreres i løpet av ett til to minutter og samtidig kunne reflektere de samme fysiske og mentale dimensjonene som i SF-36 (Campolina et al. 2011). SF-8 er således et svært tidsbesparende instrument og er lett å administrere. Instrumentet kan enten brukes som et selvrapporterende skjema eller benyttes i en form for intervju. SF-8 skårene kan sammenlignes med SF-36 skårene for de ulike subskalaene (Lefante, Jr. et al. 2005).

2.3.4 Klassifisering av pasientrapporterte utfallsmål med ICF

For å kartlegge og evaluere funksjon og funksjonshemming kan man i prinsippet benytte seg av ICF-klassifiseringen med dens 5-punkt graderte modifikator. I klassifikasjonen kan man skille mellom ulike emneområder innenfor komponentene kroppsfunksjoner (b), kroppsstrukturer (s), aktivitet og deltagelse (d,a eller p) og miljøfaktorer (e). Kodesystemet består av koder med inntil fire hierarkiske nivåer som representerer en økende grad av detaljering og presisjon (Roe et al. 2008).

Utsagnene i PGQ er klassifisert i henhold til ICF og delt inn i henhold til de to komponentene; kroppsfunksjoner og aktivitet (tabell 1). Utsagnene i ODI ble i studien til Grotle et al. (Grotle et al. 2004b) funnet å representere både kroppsfunksjoner og aktivitet/deltagelse (tabell 1). SF-8 dekker også flere komponenter (se tabell 1), men til forskjell fra ODI så summeres disse ikke i en sum skåre, men for hver enkelt skala. Generell helse-subskalaen av SF-8 kan imidlertid ikke klassifiseres innenfor de tre komponentene i ICF. Dette synliggjør en av innvendingene mot ICF, nemlig den at helserelatert livskvalitet er en dimensjon som ikke passer helt inn i modellen (Wade and Halligan 2003). Helserelatert livskvalitet er et overordnet begrep som kan dekke alle komponentene i ICF modellen samt eventuelt flere dimensjoner av helse.

Tabell 1: Klassifisering av elementene i PGQ, ODI og SF-8 med ICF

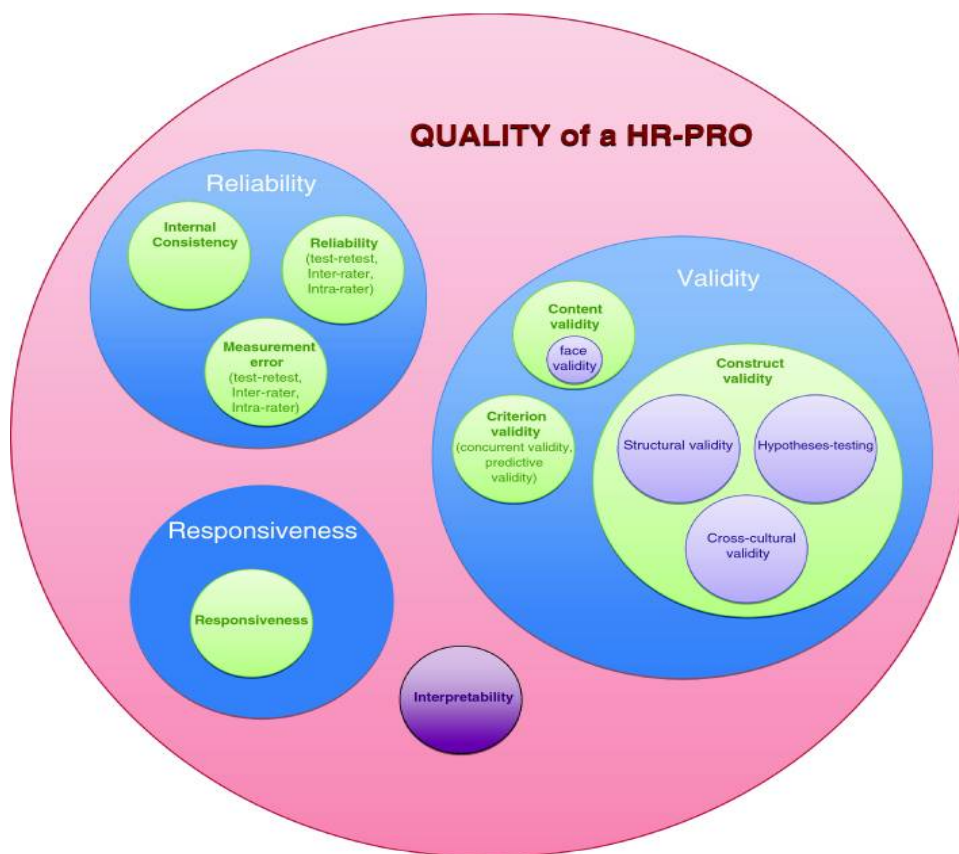
PGQ	ODI	SF8
Kroppsfunksjoner		
b2 Hvor sterke smerter har du om morgenen?	b2	b2 Kroppssmerte
b2 Hvor sterke smerter har du om kvelden?		
b7 I hvilken grad på grunn av plagene i bekkenet svikter		
b1 I hvilken grad på grunn av plagene i bekkenet gjør du		
b1 I hvilken grad på grunn av plagene i bekkenet forstyrres	b1 Sove	
		b1 Mental helse
		b1 Emosjonell rollefunksjon
Aktivitet og deltagelse		
a5 Kle på deg selv	a5 Personlig	
a4 Stå	a4 Stå	
a4 Bøye deg		
a4 Sitte	a4 Sitte	
a4 Gå	a4 Gå	a4 Fysisk funksjon
a6 Husarbeid		
a4 Bære lett		
a4 Løfte tung	a4 Løfte	
a4 Reise/sette seg		
a4 Skyve en vogn		
a4 Løpe		
a4 Utføre sportslige aktiviteter		
a4 Snu deg i sengen		
a7 Ha et normalt seksualliv	a7 Seksualliv	
a4 Skyve noe med den ene foten		
	d9 Sosialt liv	d9 Sosial funksjon
		d2 Fysisk rollefunksjon
Miljøfaktorer		
Ikke klassifiserbart i ICF		
		Generell helse
		Overskudd

b= kroppsfunksjoner, a= aktiviteter, d= deltagelse

2.4 Metode for å vurdere kvalitet på pasientrapporterte utfallsmål

Antallet tilgjengelige utfallsmål har økt dramatisk det siste tiåret. Denne økningen medfører at det kan være vanskelig å velge det beste skjemaet til det aktuelle formålet (Deyo et al. 1998; Mokkink et al. 2010; Terwee et al. 2007). Det har blitt publisert en rekke systematiske oversikter over tilgjengelige utfallsmål (Grotle et al. 2004b; Terwee et al. 2007). De systematiske oversiktene sammenligner innhold og måleegenskaper av de tilgjengelige utfallsmålene. I likhet med systematiske vurderinger av kliniske studier trenger man kriterier for å undersøke metodisk kvalitet av studier som beskriver utviklingen og evalueringen av

helsestatusskjemaer (Terwee et al. 2007). I tillegg er gode kriterier for gode måleegenskaper nødvendig for å legitimere hvilket spørreskjema som er best. Terwee (Terwee et al. 2007) har definert kvalitetskriterier for måleegenskaper for studier av utvikling og evaluering av helsestatusskjemaer. Jeg har i denne oppgaven brukt Terwees kriterier. Imidlertid har jeg delt kriteriene inn i domene; reliabilitet, validitet, responsiveness og fortolkning i henhold til COSMIN gruppens anbefaling (Mokkink et al. 2010) (se Figur 2). Som figur 2 viser så inngår flere begreper under hovedbegrepene ”reliabilitet”, ”validitet” og ”responsiveness”. I presentasjonen av hovedbegrepene vil jeg forsøke å følge COSMIN gruppens anbefalinger.



Figur 2. COSMIN gruppens anbefalinger ved testing av helse-relaterte pasientrapporterte utfallsmål. COSMIN står for Consensus-based Standards for the selection of health Measurements Instruments (Mokkink et al. 2010)

2.4.1 Reliabilitet

En av de viktigste egenskapene til alle måleinstrumenter i klinisk praksis og i forskning er at de er reliable (Devet et al. 2011). Reliabilitet er definert som ” i hvilken grad en målemetode er fri for målefeil” (Devet et al. 2011; Mokkink et al. 2010). Målinger eller registreringer vil alltid innebære en viss grad av unøyaktighet og vil ofte ikke bli oppdaget før repeterte målinger blir utført. Repeterte målinger vil vise variasjoner som oppstår fra forskjellige kilder; selve måleinstrumentet, personer som utfører testen, pasienten som blir testet eller under hvilke forhold testen er utført (Devet et al. 2011). Reliabiliteten står på spill i alle disse variasjonene som oppstår under målingene (Devet et al. 2011). For eksempel kan tidsperioden mellom de repeterte målingene påvirke reliabiliteten av et instrument. Tiden bør være lang nok for å hindre at test personene husker hva de svarte første gangen, men perioden må være kort nok til at det ikke skjer forandringer i mellomtiden. Ofte er 1-2 uker hensiktsmessig, men det kan være grunner til å velge annerledes.

Ulike termer blir ofte tatt i bruk for å beskrive reliabilitet; reproduserbarhet, enighet, repeterbarhet og samsvar mellom målinger, men få av dem har en formell definisjon (Steiner 2008). COSMIN gruppen (Mokkink et al. 2010) har delt begrepet reliabilitet inn i; intern konsistens, reliabilitet og målefeil (Figur 2).

Intern konsistens (internal consistency) omhandler i hvor stor grad utsagnene i et instrument er relatert til hverandre slik at de måler det samme begrepet eller fenomenet (Terwee et al. 2007). Intern konsistens er en viktig egenskap for instrumenter som er ute etter å måle et underliggende fenomen. Hvis man for eksempel har til hensikt å måle aktivitet i dagliglivet må alle spørsmålene kunne relateres til dette. I klassisk måleteori brukes faktoranalyse og Cronbach's alpha for å undersøke om utsagnene i spørreskjemaene måler det samme konstruktet (Terwee et al. 2007). En Cronbach's alpha mellom 0.70 og 0.95 regnes som god intern konsistens.

Reliabilitet henspeiler i hvilken grad pasienter kan bli atskilt fra hverandre til tross for målefeil (Terwee et al. 2007). Høy reliabilitet er viktig når man ønsker å skille mellom pasienter med høy eller lav grad av det målte fenomenet, for eksempel høy eller lav grad av aktivitetsbegrensninger i dagliglivet. Intraclass correlation coefficient (ICC) er den vanligste metoden å uttrykke reliabilitet på når man har instrumenter med kontinuerlige totalskårer. Terwee gruppen skiller mellom ICC for konsistens og ICC for agreement. Fordi ICC for

agreement tar høyde for systematiske forskjeller mellom to eller flere målinger foretrekkes denne. Pearson korrelasjons koeffisient er ikke tilstrekkelig mål på reliabilitet da denne ikke tar høyde for systematiske forskjeller (Terwee et al. 2007). Ved ordinale eller kategoriske totalskårer så benyttes fortrinnsvis en vektet Kappa koeffisient for å uttrykke reliabiliteten av to eller flere målinger (Terwee et al. 2007). En ICC eller Kappa koeffisient på 0.70 eller mer regnes som god reliabilitet.

Terwee omtaler målefeil for *agreement* og det peker på avstanden mellom repeterte målinger gitt at tilstanden er den samme. Dette uttrykkes ofte som *standard error of measurement* (SEM), evt. som *smallest detectable change* (SDC) når man tar høyde for 95 % konfidens intervaller rundt målefeilen (Terwee et al. 2007).

2.4.2 Validitet

Forholdet mellom reliabilitet og validitet er unidireksjonelt. Det betyr at reliabilitet er en forutsetning, men ingen garanti for validitet (Portney and Watkins 2000). Dersom et måleinstrument ikke produserer stabile målinger under like forhold, med andre ord er reliabelt, kan det heller ikke anses å være valid (Batterham and George 2003). Validitet er definert av COSMIN som: ”i hvilken grad en målemetode virkelig måler det den har i hensikt å måle” (Devet et al. 2011; Mokkink et al. 2010). Denne definisjonen kan se enkel og grei ut, men det har vært mye diskusjon om hvordan validitet skal testes og hvordan resultatene bør tolkes (Devet et al. 2011). Man skiller ofte mellom måleinstrumentets innholdsvaliditet, kriterievaliditet, begrepsvaliditet og diskriminerende validitet (Devet et al. 2011; Mokkink et al. 2010; Steiner et al. 2002) (Figur 2).

Innholdsvaliditet (content validity) omhandler i hvilken grad det man ønsker å måle blir fanget opp av utsagnene i et funksjonsmåleinstrument (Terwee et al. 2007). Innholdsvaliditet krever at et instrument ikke inneholder faktorer som er irrelevant for hensikten med instrumentet (Portney and Watkins 2000). Undersøkelse av innholdsvaliditet er i hovedsak en subjektiv prosess, der man forsøker å oppnå konsensus i en gruppe bestående av forskere, klinikere og eventuelt pasientrepresentanter. Terwee (Terwee et al. 2007) har satt opp følgende kriterier for å vurdere innholdsvaliditet;

Målet med instrumentet. Det er viktig å beskrive hvilket fenomen eller hovedbegrep instrumentet har til hensikt å måle, for eksempel smerte, funksjon osv.

Målepopulasjon. Det er viktig å beskrive hvem instrumentet er utviklet for.

Hva instrumentet har til hensikt å måle. Det er viktig at forfatterne klart beskriver hvilket fenomen eller hovedkonsept som skal måles.

Utvelgelse av utsagn (item). Metoden for hvordan man går frem når man utvikler og selekterer utsagn, spørsmål eller såkalte item må beskrives inngående. Videre bør man beskrive om det er gjort pilot studier for å undersøke lesbarheten og forståelsen av et spørreskjema, eventuelt testing hvis det er et testbatteri. Ved utvikling av selvrappporterte spørreskjema starter man ofte med et stort antall utsagn, som deretter reduseres i løpet av uttestingen av instrumentet.

Fortolkning av utsagnene. Det skal ikke kreve mer lese- skriveferdigheter enn det en 12 åring er i stand til, dette for å unngå manglende svar ("missing") og upålitelige svar. Det betyr at utsagnene skal være korte og enkle og skal ikke inneholde vanskelige ord og sjargonger. Ett utsagn skal ikke inneholde to spørsmål, videre bør tidsintervallet som spørsmålet refererer til være klart og tydelig.

Kriterium validitet (criterion validity) refererer til i hvilken grad et utfallsmål er relatert til en gullstandard (Mokkink et al. 2010). Det innebærer at man ikke kan vurdere kriterium validitet hvis det ikke er noen gullstandard tilgjengelig. Terwee (Terwee et al. 2007) gir positiv skåring hvis det er overbevisende argumenter for at de utfallsmålene som brukes som "gullstandard" er relevante og at korrelasjonen er minst 0.70.

Begrepsvaliditet (construct validitet) blir definert av COSMIN panelet som; "i hvilken grad resultatene av måleinstrumentet er konsekvent med hypotesene med hensyn til interne forhold, relasjoner med en rekke andre instrumenter eller forskjeller mellom relevante grupper" (Mokkink et al. 2010). Hypotesene kan være forventede høye korrelasjoner mellom instrumenter som måler det samme begrepet (constructet), eventuelt lave til moderate korrelasjoner mellom instrumenter som måler noe ulike begreper, for eksempel smerte og aktivitetsbegrensninger hos bekkenpasienter. Hypotesene kan også dreie seg om forventede forskjeller i skåringene mellom "kjente" grupper, for eksempel mellom gravide og ikke-gravide kvinner med bekkenrelaterte plager (Terwee et al. 2007). Når man undersøker forventede forskjeller mellom "kjente" subgrupper i et utvalg blir det omtalt som diskriminerende validitet (Devet et al. 2011). Terwee gir positiv skåring for begrepsvaliditet

hvis hypoteser vedrørende relasjonene til andre spørreskjemaer som måler samme underliggende begrep er spesifisert i forkant og minst 75 % av resultatene korresponderer med disse hypotesene.

COSMIN gruppen anbefaler også at begrepsvaliditet bør vurderes ved å teste apriori hypoteser, som for eksempel forventet korrelasjon mellom forandring i målinger, eller forventede forskjeller i forandringer mellom kjente grupper.

En nyere og bedre måte å undersøke begrepsvaliditet på er å anvende såkalt item - response teori og Rasch analyse, noe som ble anvendt i utviklingen og testingen av PGQ (Stuge et al. 2011). Det faller utenfor denne oppgavens rammer å gi en inngående beskrivelse av denne tilnærmingen, men en av hovedanalysene i Rasch analysen er å undersøke hvorvidt et instrument er unidimensjonalt, dvs at det måler ett "construct" (fenomen) og ikke flere. Videre får man gjennom Rasch analyse en god undersøkelse av om måleskalaen fungerer optimalt for det målte fenomenet.

I tillegg til disse nevnte undersøkelsene av reliabilitet og validitet er det viktig å undersøke om det er en tak - eller gulv effekt i skåringene. Tak eller gulv effekt kan oppstå når en høy andel av den totale populasjonen har skåringer på den laveste eller øverste delen av skalaen (Devet et al. 2011). Tak eller gulveffekt blir vurdert hvis mer enn 15 % av respondentene skårer på den laveste eller høyeste skåren. Terwee gir positiv vurdering hvis det ikke er tak - eller gulv effekt.

2.4.3 Responsiveness – diskriminerende longitudinell validitet

Responsiveness er en av de mest sentrale kvalitetene ved et utfallsmål der hensikten er å måle en endring i klinisk status når man forventer at det har skjedd en endring. Det betyr at en skåring må forandre seg i takt med at pasientens status forandres eller den må være uforandret hvis pasienten er uforandret (Portney and Watkins 2000). Responsiveness blir definert som et utfallsmåls evne til å fange opp endring over tid når en endring har skjedd (Devet et al. 2011). Dette kan også uttrykkes som longitudinal validitet.

Mange metoder har blitt benyttet for å undersøke responsiveness av et instrument (Terwee et al. 2007). Den vanligste metoden i mange år har vært å se på ulike typer effektstørrelse, slik som for eksempel Grotle og medarbeidere gjorde i sin studie fra 2004 (Grotle et al. 2004a) .

Fordi responsiveness handler om å kunne diskriminere mellom viktig klinisk endring eller ikke så har Receiver Operating Characteristic (ROC) kurver blitt mer og mer benyttet, og regnes som et bedre mål på responsiveness enn effektstørrelser. Terwee og COSMIN gruppen anerkjenner ikke effektstørrelser som gyldige mål på responsiveness, men de anbefaler ROC kurve analyser. Ved å benytte seg av ROC kurve analyser så kan man få estimerer for minste klinisk viktige endring (minimal important change - MIC). Hvis denne verdien er større enn instrumentets målefeil (SDC), så angir Terwee positiv vurdering. Videre hvis arealet under ROC kurven er større eller lik 0.70, så har også instrumentet tilstrekkelig diskriminerende evne og responsiveness (Terwee et al. 2007).

2.4.4 Fortolkning

Fortolkning (interpretability) er definert som ” i hvilken grad man kan tillegge kvalitativ mening til kvantitative skår” (Mokkink et al. 2010). Forskeren bør sørge for at det er informasjon om hva som gir klinisk mening. Man kan få klinisk mening av enkelt skåringer og endring i skåringer (Devet et al. 2011).

2.5 Hoved- og delproblemstillinger

Hovedproblemstillingen i denne masteroppgaven er å undersøke diskriminerende validitet av de tre utfallsmålene; PGQ, ODI og SF-8 når det benyttes hos kvinner med eller uten bekkenleddssmerter etter fødsel.

Hypotese: Det var forventet at kvinner med bekkenleddssmerter skulle skåre høyere på alle utfallsmålene, spesielt på PGQ og ODI som er to sykdomsspesifikke instrumenter, enn kvinner uten bekkenleddssmerter. Det var forventet en mindre forskjell i skåringene på det generiske instrumentet SF-8, spesielt på den mentale komponent skåren.

Del- problemstillinger

Den første del- problemstillingen omhandler følgende spørsmål: Viser skåringene i PGQ, ODI og SF-8 en normalfordeling hos kvinner med og uten bekkenleddssmerter? Den andre var: Er det tak- eller gulveffekt i skåringene på de tre instrumentene når de brukes hos kvinner med og uten bekkenleddssmerter.

Hypoteser: Det var forventet at skåringene hos kvinner med bekkenleddssmerter var normalfordelte i alle instrumentene. Det var forventet at skåringene hos kvinner uten bekkenleddssmerter ville være svært lave og ikke normalfordelte i alle instrumentene. Hos kvinner uten bekkenleddssmerter var det forventet en gulveffekt i skåringene på PGQ og ODI fordi dette er to sykdomsspesifikke instrumenter. Det var ikke forventet gulv- eller takeffekt på SF-8 hos kvinner uten bekkenleddssmerter fordi dette er et generisk instrument.

3 Metode

3.1 Studiedesign

Denne masteroppgaven er en tverrsnittstudie av kvinner med og uten bekkenleddssmerter. Alle kvinnene gjennomgikk en klinisk testing før inklusjon. De svarte så på et spørreskjema som blant annet inneholdt sosiodemografisk bakgrunnsinformasjon og instrumentene PGQ, ODI og SF-8.

3.2 Materiale

3.2.1 Bakgrunn for valg av deltagere

Som nevnt i innledningen er denne masteroppgaven en del studie av bekkenprosjektet "Bekkenleddsplager og bekkenbunn - noen sammenheng?" Da jeg var ansvarlig for inkludering og klinisk testing av pasientene med bekkenleddssmerter til studien hvor vi utviklet PGQ, var det naturlig også å bidra til rekruttering av pasienter i det ovenfor nevnte prosjektet. Dette prosjektets materiale med 50 kvinner med og 50 kvinner uten bekkenleddssmerter egnet seg godt til å kunne besvare mine problemstillinger for denne masteroppgaven.

3.2.2 Inklusjonskriterier for kvinner med bekkenleddssmerter

Kvinnene skulle ha bekkenleddssmerter av minimum 6 måneders varighet lokalisert distalt og/eller lateralt for L5-S1 området, i setet og/eller i symfysen oppstått under svangerskapet eller innen 3 uker etter fødsel. Diagnostiseringen av bekkenleddssmerter ble videre basert på følgende tester; Posterior Pelvic Pain Provocation (P4) test, Active Straight Leg Raise (ASLR) test, Distraksjonstest, Kompresjonstest, Patrick-Fabers test, Test av Long Dorsal Ligament, Smerteprovokasjon ved palpasjon av symfysen og Modifisert Trendelenburgs test (Stuge et al. 2004a; Vleeming et al. 2008). For å bli diagnostisert til å ha bekkenleddssmerter måtte pasienten ha positiv P4 test og/eller ASLR test (skår over 3) og minst en positiv av de tre andre testene. Kvinnene måtte beherske norsk både skriftlig og muntlig.

Beskrivelse av kliniske tester for bekkenleddssmerter

Smerteprovokasjon av bakre bekken (P4-testen)

P4-testen ble utført med pasienten liggende på ryggen og hoften på den siden som ble testet flektert til 90 grader og kneet maksimalt flektert. Det andre benet lå strakt. Undersøker ga et lett trykk i femurs lengderetning mens hun stabiliserte bekkenet med den andre hånden på pasientens motsatte spina iliaca anterior superior. Testen var positiv når pasienten kjente en gjenkjennende, lokalisert smerte i iliosakralleddet på den testede siden (Ostgaard et al. 1994).

Distraksjons test

Distraksjons test ble utført med pasienten i ryggliggende stilling. Undersøker plasserte armene i kryss og ga press på anterior superior iliaca anterior superior i lateral retning. Testen var positiv når den provoserte frem kjente smerter i bakre del av bekkenet (Laslett et al. 2005b).

Kompresjons test

Kompresjons test ble utført med pasienten i ryggliggende med strake ben. Undersøker plasserte hendene lateralt for anterior superior iliaca superior og komprimerte i medial retning bilateralt. Testen var positiv når den provoserte frem smerter i bakre del av bekkenet (Kristiansson and Svardsudd 1996).

Patrick-Faber test

Patrick-Faber testen ble utført med pasienten liggende på ryggen med det ene benet i fleksjon, abduksjon og utadrotasjon slik at helen hvilte på det motsatte kneet. Undersøker påførte et lett trykk over kneet med den ene hånden mens den andre kontrollerte den motsatte hoften. Testen var positiv når pasienten kjente smerter bak i bekkenet og/eller når pasienten hadde nedsatt bevegelse (Broadhurst and Bond 1998; Strender et al. 1997).

Test av de lange dorsale iliosakralleddsligamentene

Pasienten lå på magen. Undersøker palperte de lange dorsale iliosakralleddsligamentene direkte under kaudale del av spina iliaca posterior superior. Testen var positiv hvis

palpasjonen provoserte frem smerter. Begge sidene ble testet og skåret separat (Vleeming et al. 2002).

Smerte provokasjon av symfysis pubis ved palpasjon

Palpasjon av symfysis pubis ble utført med pasienten liggende på ryggen. Undersøker påførte et skånsomt trykk over symfysis pubis med flate fingre. Testen var positiv hvis smerten påført av trykket vedvarte mer enn fem sekunder etter at fingrene ble fjernet (Albert et al. 2002).

Smerteprovokasjon av symfysis pubis ved modifisert Tredelenburgs test.

Pasienten ble bedt om å stå på et ben og løfte det andre benet opp til 90 grader i hoften og kneet. Hvis det å stå på ett ben provoserte frem smerter i symfysis pubis ble testen skåret som positiv. Testen ble utført på begge sider (Albert et al. 2002).

Aktiv strakt benløft (ASLR) test

ASLR testen ble utført med pasienten liggende på ryggen med strake ben ca. 20 cm fra hverandre. Pasienten løftet strakt ben 20 cm over benken uten å bøye kneet og ble bedt om å skåre hvor tungt eller vanskelig det var å løfte benet på en seks punkts skala fra 0 "ingen problemer" til 5 "umulig å løfte". Begge sider ble testet og skåringene på begge sider ble summert slik at man fikk en sum skår fra 0 til 10 (Mens et al. 2001).

3.2.3 Inklusjonskriterier for kvinner uten bekkenleddssmerter

Gruppen med kvinner uten bekkenleddssmerter skulle ikke ha bekkenleddssmerter, rygg- eller underlivsplager. Kvinner uten bekkenleddssmerter ble for øvrig matchet med kvinnene med bekkenleddssmerter hva gjelder alder (± 5 år), antall vaginale fødsler og antall barn. De inkluderte pasientene som hadde barn nede i 6 måneders alderen ble matchet med alle barn under 1 år ± 1 mnd, og alle barn over 1 år til ± 5 år. Dette ble gjort for at gruppene skulle bli likest mulig med tanke på endringer i bekkenbunn etter fødsel. For å bli inkludert i studien måtte alle beherske norsk både skriftlig og muntlig.

3.2.4 Eksklusjonskriterier for begge gruppene

Kvinner som hadde født barn for mindre enn 6 måneder siden, fortsatt ammet, hadde korsryggsmerter med radierende utfall, nevrologiske sykdommer, kreft, psykiatriske lidelser, ubehandlet urinveisinfeksjon og tidligere kirurgi av bekkenbunnen ble ekskludert fra studien

3.3 Datainnsamling

Kvinnene som samtykket til å delta i studien "Bekkenleddsplager og bekkenbunn – noen sammenheng?" ble testet i henhold til inklusjonskriteriene, før de fylte ut spørreskjema som inneholdt sosiodemografiske data (alder, utdanning, arbeidsstatus, antall barn og vaginale fødsler, høyde og vekt), generell helsestatus SF-8 (Ware, Jr. and Sherbourne 1992), smerte og funksjonsstatus målt med PGQ (Stuge et al. 2012) og ODI (Fairbank et al. 1980). Alle kvinnene ble klinisk testet i løpet av en 8 ukers periode av fysioterapeut Kaja Sætre, som var prosjektleder for studien "Bekkenleddsplager og bekkenbunn – noen sammenheng?".

I første fase av rekrutteringen av pasienter ble det opprettet kontakt med fysioterapeuter og institutter som hadde erfaring og tilgang til kvinner med bekkenleddssmerter. Institutter som ble kontaktet i denne prosessen var; Fagerborg fysioterapi, Kiropraktor klinikken Bekken og barn, Føyka Fysioterapi, Galgeberg fysikalske, Gruppefysio AS, Ullernklinikken, Sentralbanestasjonen Fysikalske, Hans og Olaf Fysioterapi og Hjelp24NIMI, Ullevål. Aktuelle behandlere ble oppsøkt og informert om prosjektet. I tillegg ble pasienter rekruttert direkte fra gruppetreninger for kvinner med bekkenleddssmerter.

3.3.1 Inkludering av kvinner med bekkenleddssmerter

Det ble holdt møter med behandlere for å oppmuntre i forhold til rekruttering av kvinner med bekkenleddssmerter og i tillegg ble det hengt opp skriftlig informasjon på instituttene og enkelte helsestasjoner i Oslo området. Videre markedsføring ble gjort ved at informasjon ble lagt ut på hjemmesiden til Landsforeningen for kvinner med bekkenløsningssmerter, avdeling Oslo. Facebook ble også brukt og prosjektet ble promotert på ulike "mammablogger". De som var interesserte tok kontakt med prosjektleder og fysioterapeut Kaja Sætre direkte på e-post eller telefon og fikk både muntlig og skriftlig informasjon om prosjektet. Deltakerne som sa seg villig til å delta i prosjektet fikk avtalt tid til testing på Hjelp24NIMI. Inklusjonstestingene varte 60 minutter og gikk med til å teste pasienten etter protokollen og utfylling av

spørreskjemaer. Det ble gitt anledning for kvinnene å stille spørsmål de måtte ha med tanke på egen situasjon og i forhold til prosjektet.

3.3.2 Inkludering av kvinner uten bekkenleddssmerter

Inkludering av kvinner uten bekkenleddssmerter foregikk stort sett gjennom venner og bekjente av de inkluderte kvinnene med bekkenleddssmerter. Det ble i tillegg markedsført via Facebook der svært mange meldte sin interesse. Alle interesserte tok direkte kontakt med prosjektleder Kaja Sætre på telefon eller e-post og mottok nøyaktig samme informasjon som pasientene både skriftlig og muntlig. De fikk tilsendt spørreskjemaene i ferdig frankerte svarkonvolutter som ble returnert til prosjektleder.

3.4 Pasientrapporterte utfallsmål

3.4.1 Pelvic Girdle Questionnaire (PGQ)

PGQ består av 20 funksjonsspørsmål og fem symptomspørsmål, som skåres på en firepunktsskala. Funksjonsspørsmålene skåres på en firepunkts deskriptiv skala; ikke i det hele tatt = 0, i liten grad = 1, i noen grad = 2 og i stor grad = 3. Symptomspørsmålene blir også skåret på en firepunktsskala. Smerter ble gradert slik; ingen=0, noe=1, moderate=2 og svært mye=3, de øvrige 3 symptomspørsmålene blir gradert; ikke i det hele tatt=0, i liten grad=1, i noen grad =2 og i stor grad =3. Sum skårene blir summert og regnet ut i prosent fra 0 – 100 hvor 100 er verst mulig skår (Grotle et al. 2012; Stuge et al. 2011).

PGQ har vist god intern konsistens, test-retest reliabilitet, og begrepsvaliditet i en tidligere studie (Stuge et al. 2011; Grotle et al. 2012). Videre har PGQ funnet å kunne diskriminere mellom gravide og ikke-gravide kvinners skåringer samt skåringer hos kvinner med og uten smerte i flere områder i bekkenregionen.

3.4.2 Oswestry Disability Index, versjon 2.0 (ODI 2.0)

Den anbefalte versjon 2.0 av ODI, som har blitt oversatt og testet for norske pasienter (Grotle et al. 2003) ble benyttet i vår studie. Denne versjonen inneholder 10 spørsmål som er relatert til daglige aktiviteter og/eller smerte (smerteintensitet, personlig stell, løfte, gå, sitte, stå, sove, seksualliv, sosialt liv og reising). Hvert spørsmål inneholder seks utsagn som beskriver i en økende grad vanskelighet relatert til gjeldene aktiviteten. De blir skåret på en 6 punkts skala (0-5) hvor 0 representerer ingen nedsatt funksjon og 5 representerer maks nedsatt funksjonsevne. Skåringene blir summert og prosentuert til en 0-100 sum skåre der 0 er ingen funksjonsproblem og 100 er maksimale funksjonsproblemer (Baker et al. 1990; Grotle et al. 2003). Alvorlighetsgraden av funksjonsproblemer blir ofte inndelt i følgende kategorier; 0-20 minimale funksjonsproblemer, 20-40 moderate funksjonsproblemer, 40-60 alvorlige funksjonsproblemer, 60-80 invalidiserende funksjonsproblemer, 80-100 sengeliggende (Fairbank et al. 1980; Resnik and Dobrykowski 2005).

Den norske ODI versjonen ble oversatt og tilbakeoversatt i henhold til Guillemín kriteriene (Guillemín F. et al. 1993) av Grotle og medarbeidere (2003). Skjemaet ble testet ut på 55 pasienter med akutte korsryggsmarter, samt 50 pasienter med kroniske korsryggsmarter. Den norske versjonen av ODI viste i denne studien en god intern konsistens med Cronbachs alpha på 0.94, god test-retest reliabilitet med en ICC på 0.88, samt en målefeil på 12%.

Begrepsvaliditeten var også som forventet da ODI 2.0 versjonen korrelerte høyt med fysisk funksjonsskalaen av SF-36, moderat med smerteskalaen og lavt med mentalskalaen av SF-36. Grotle (2004a) undersøkte også responsiveness av denne ODI versjonen og fant at den hadde god responsiveness hos både gruppen med akutte og kroniske korsryggsmarter.

3.4.3 SF- 8 Health Survey (SF- 8)

SF-8 er som tidligere nevnt en kortversjon av SF-36 og inneholder ett spørsmål angående generell helse (SF-1), fysisk funksjon (SF-2), fysisk rollefunksjon (SF-3), kropps smerte (SF-4), overskudd (SF-5), sosial funksjon (SF-6), mental helse (SF-7) og emosjonell rollefunksjon (SF-8). Spørsmålene skåres på en ordinal skala fra 0 til 4 eller 5 der 0 er best skåre og 4 eller 5 dårligst skåre. Skårene summeres og omregnes til en prosentuert skåre fra 0 til 100, der 0 er dårlig helse og 100 er maksimalt god helse. Som i SF-36 benyttes en syntax i beregning av de prosentuerte SF-8 skårene slik at skåringene fra 0 til 100 er vektet for fordeling i

normalbefolkningen. En skåre over eller under 50 regnes som over eller under gjennomsnittet i den generelle befolkningen i USA (Ware et al. 2001).

I denne oppgaven presenteres både råskårene i alle delskalaene av SF-8, og de summerte fysiske og mentale komponent skårene. Fysisk komponent skåren (PCS) ble beregnet ut i fra en summering av fysisk funksjon, fysisk rollefunksjon, og kroppssmerte, mens mental komponent skåren (MCS) ble beregnet ut ifra en summering av mental helse, emosjonell rollefunksjon og sosial funksjon (Ware et al. 2001).

Test-retest reliabilitet av SF-8 viste seg å være svært god i den norske bekkenstudien der SF-8 ble sammenlignet med PGQ (Grotle et al. 2012). ICC var akseptabel til god og varierte mellom 0.78 og 0.89, og målefeil ble estimert til mellom 7% og 13% av totalskårene.

Testingen av begrepsvaliditeten bekreftet at SF-8 skalaene måler noe annet enn funksjon- og symptomskalaene av PGQ, samt også ODI og DRI. Flere av SF-8 skalaene diskriminerte imidlertid ikke tilfredsstillende mellom gravide og ikke-gravide kvinner og kvinner med bekkenleddssmerter i alle tre bekkenleddene og kvinner med bekkenleddssmerter i ett eller to ledd (Grotle et al. 2012). Ett unntak var fysisk funksjon delskalen i SF-8 som diskriminerte mellom de nevnte subgruppene i Grotles studie (2012).

3.5 Databehandling

Spørreskjemaene ble lagret i innelåst skap på OUS. Dataene ble plottet manuelt i SPSS 16.0 og lagret på serveren "Vilje" ved OUS.

De demografiske variablene (alder, utdanning, arbeidsstatus, antall barn) ble analysert med frekvensanalyser. Frekvensanalyser og eksplorerende statistikk ble også gjort for hvert enkelt utsagn i de tre utfallsmålene, PGQ, ODI og SF-8, for å undersøke kvaliteten på dataene i form av missing verdier, gjennomsnitt og median av råskårene, samt laveste og høyeste skår på hvert enkelt utsagn. Normalfordelingen av del- og totalskårer i de tre instrumentene, PGQ, ODI og SF-8 ble videre analysert i histogramplott. Disse analysene ble gjort separat for kvinner med og uten bekkenleddssmerter. Del- og totalskårer i de tre instrumentene ble videre presentert i boxplott, også separat for de to gruppene i studien.

Tak- og gulveffekt ble undersøkt ved å observere andelen av personer som skåret høyest eller laveste skåre på instrumentene. Hvis mer enn 15 % av respondentene skåret på den laveste eller høyeste skåren ble dette angitt som tak - eller gulveffekt.

For å teste hypotesen om at kvinner med bekkenleddssmerter ville skåre høyere på de tre instrumentene sammenlignet med kvinner uten slike plager, ble den ikke-parametriske t-testen Mann-Whitney test benyttet. Denne testen ble benyttet fordi materialet ikke var normalfordelt. Istedenfor å sammenligne gjennomsnittet i de to gruppene som er vanlig i en t-test så sammenligner Mann-Whitney testen medianen i gruppene (Altman 1991). For å vurdere diskriminerende validitet ble Receiver operating characteristic ROC analyser. Gruppe-variablen (kvinner med og uten bekkenleddssmerter) ble brukt som eksternt kriterium i ROC analysen, og kurvene gjenspeilte derfor instrumentenes sensitivitet (y-aksen) og spesifisitet (x-aksen) til å skille mellom de to gruppene. Arealet under ROC kurven gir et estimat på instrumentenes nøyaktighet til å skille mellom to grupper. Hvis instrumentet skiller helt perfekt vil arealet under kurven ha en verdi på 1, mens en verdi på 0.5 betyr at det er 50 % sjanse for å skille mellom to grupper (Altman 1991; Grotle et al. 2012).

3.6 Etisk godkjenning

Studien er godkjent av Regional komité for forskningsetikk. Før kvinnene ble klinisk testet og før de fylte ut skjemaene, signerte de en samtykkeerklæring med informasjon om at de kunne trekke seg når de måtte ønske det uten å måtte oppgi noen grunn. Informasjon og data ble anonymisert og oppbevart utilgjengelig for utenforstående på forskningsserver ved Oslo Universitetssykehus.

4 Resultater

4.1 Utvalget

Totalt 49 kvinner med og 49 kvinner uten bekkenleddssmerter bel inkludert i studien.

Kvinnene hadde en gjennomsnittsalder på 35 år i begge gruppene. Gjennomsnittsalderen var 35 år i begge gruppene, og det var høy andel av høy utdanning i begge gruppene. Ingen av kvinnene uten bekkenleddssmerter var sykemeldte, mottok rehabiliteringspenger eller uføretrygd, mens det var 9 av kvinnene med bekkenleddssmerter som var sykemeldte, 3 mottok rehabiliterings penger og 3 var uføretrygdet. Som vist i tabell 2 hadde kvinnene med bekkenleddssmerter i gjennomsnitt 1.8 barn, mens de friske hadde 1.7. Femti-tre prosent av kvinnene med bekkenleddssmerter hadde smertefrie perioder av og til, mens 89 % av de friske kvinnene hadde ofte smertefrie perioder. Kvinnene med bekkenleddssmerter i gjennomsnitt hatt smerter i 3.4 år (Tabell 2).

Tabell 2: Demografiske data av kvinner med og uten bekkenleddssmerter.

	Kvinner uten bekkenleddssmerter N=49	Kvinner med bekkenleddssmerter N=49
Alder år, mean (SD) median	35 (4.8) 34 (18)	35 (4.8) 35 (21)
Utdanning, antall år >12 år n (%) mean median (IQR)	46 (93.9) 16.7 16 (10.0)	43 (87.8) 16.1 17 (17.5)
Arbeids status, n (%) [*] Ansatt - i jobb - sykemeldt - fødselspermisjon Rehabilitering Uføretrygdet Student Annet	 33 (67) 0 14 (28.6) 0 0 2 (4.1) 0	 11 (22.4) 9 (18.4) 15 (30.6) 3 (6.1) 3 (6.1) 6 (12.2) 1 (2.0)
Antall barn, n (%) 1 2 3 4	 24 (49.0) 19 (38.8) 4 (8.2) 2 (4.1)	 21 (43) 20 (39) 6 (12) 2 (4)
Smertefrie perioder, n (%) - aldri - av og til - ofte	 1 (2) 4 (8.2) 44 (89.8)	 6 (12) 26 (53) 17 (35)
Smertevarighet, år mean (SD) Median (IQR)	 0.27 (1.4) 0 (10)	 3.4 (3) 2 (13.7)

* = 1 missing

4.2 Datakvalitet

Tabell 3 og 4 presenterer datakvaliteten i de tre instrumentene for henholdsvis kvinner med og uten bekkenleddssmerter. Det var ingen missing verdier i skåringene hos kvinnene med bekkenleddssmerter, og kun en missing på SF-8 i gruppen med kvinner uten bekkenleddssmerter.

Gjennomsnitt- og medianverdiene i tabell 3 er presentert med laveste verdi på toppen og økende verdier nedover i tabellen. På denne måten anskueliggjør tabellen hvilket utsagn kvinnene med bekkenleddssmerter skårer lavest på (dvs det letteste utsagnet der de har minst problemer) og de utsagn de skårer høyest på (dvs de utsagnene de har mest problemer med). For eksempel, på PGQ Aktivitet skårer kvinnene lavest på utsagn 7 (*Gå mindre enn 10 minutter*) og høyest på utsagn 12 (*Løfte tungt*). På ODI skårer kvinnene lavest på utsagn 2 (*Personlig stell*) og høyest på utsagn 6 (*Stå*). På SF-8 skalaene skårer kvinnene lavest på emosjonell rollefunksjon og høyest på kroppssmerte skalaen.

Som tabell 4 viser så skårer kvinnene uten bekkenleddssmerter mye lavere på alle utsagnene i alle instrumentene enn kvinnene med bekkenleddssmerter. På PGQ og ODI skårer de lavt på samtlige utsagn, mens på SF-8 skårer de noe høyere. De skårer lavest på fysisk rollefunksjon og høyest på overskudd skalaen.

Tabell 3: Skåringer på hvert enkelt utsagn i de tre instrumentene – PGQ, ODI, og SF-8- for kvinner med bekkenleddssmerter; missing data, gjennomsnitt skår med standard avvik (SD) median med interkvartil range (IQR), andeler med laveste og høyeste skår

n=49	Missing data	Mean (SD)	Median (IQR)	Lavest skår %	Høyest skår %
PGQ aktivitet(0-3)					
7. Gå mindre enn 10 minutter	0	0.55 (0.81)	.00 (2)	61.2	0
5. Sitte mindre enn 10 minutter	0	0.61 (0.81)	.00 (3)	55.1	4.1
2. Stå mindre enn 10 minutter	0	0.71 (0.82)	.00 (2)	51.0	0
11. Bære lett	0	0.84 (0.85)	1.00 (3)	42.9	2
1. Kle på deg selv	0	0.84 (0.94)	1.00 (3)	46.9	6.1
17. Ligge	0	1.16 (0.97)	1.00 (3)	30.6	8.2
18. Snu deg i sengen	0	1.18 (1.11)	1.00 (3)	36.7	16.3
13. Reise/sette seg	0	1.27 (0.88)	1.00 (3)	18.4	10.2
19. Ha et normalt seksualliv	0	1.31 (0.92)	1.00 (3)	20.4	10.2
4. Bøye deg	0	1.57 (1.0)	2.00 (3)	18.4	18.4
9. Gå trapper	0	1.59 (1.04)	2.00 (3)	18.4	20.4
14. Skyve en vogn	0	1.76 (1.01)	2.00 (3)	14.3	26
20. Skyve noe med den ene foten	0	1.82 (1.05)	2.00 (3)	16.3	30.6
10. Husarbeide	0	1.86 (0.91)	2.00 (3)	10.2	24.5
6. Sitte mer enn 60 minutter	0	1.98 (1.11)	2.00 (3)	14.3	44.9
8. Gå mer enn 60 minutter	0	2.14 (1.04)	2.00 (3)	12.2	49
3. Stå mer enn 60 minutter	0	2.31 (0.91)	3.00(3)	8.2	53.1
16. Utføre sportslike aktiviteter	0	2.35 (0.86)	3.00 (3)	4.1	55.1
15. Løpe	0	2.45 (0.91)	3.00 (3)	6.1	67.3
12. Løfte tungt	0	2.51(0.82)	3.00 (3)	4.1	67.3
PGQ symptom (0-3)					
23. Svikter benet under deg?	0	0.90 (1.09)	.00 (3)	51	12.2
21. Smerter om morgenen	0	1.29 (0.84)	1.00 (3)	16.3	8.2
25. Forstyrres nattesøvnen?	0	1.31 (1.03)	1.00 (3)	24.5	16.3
24. Gjør du ting langsommere?	0	2.00 (0.96)	2.00 (3)	8.2	36.7
22. Smerter om kvelden	0	2.12 (0.86)	2.00 (3)	4.1	38.8
ODI (0-5)					
2. Personlig stell	0	0.43 (0.74)	.00 (2)	71.4	0
4. Gå	0	1.2 (0.97)	1.00 (4)	24.4	0
7. Sove	0	1.6 (0.90)	1.00 (4)	18.4	0
3. Løfte	0	1.8 (1.17)	1.00 (4)	2.0	0
10. Reising	0	1.27 (1.00)	1.00 (4)	16.3	0
8. Seksualliv	0	1.27 (1.20)	1.00 (4)	26.5	0
9. Sosialt liv	0	1.88 (1.07)	1.00 (4)	16.3	0
5. Sitte	0	1.98 (1.13)	2.00 (4)	12.2	0
1. Smerteintensitet	0	2.29 (0.98)	2.00 (3)	20.4	0
6. Stå	0	2.33 (1.23)	2.00 (4)	8.2	0
SF-8					
SF-8 Emosjonell rollefunksjon (0-4)	0	1.43 (1.31)	1.00 (5)	30.6	8.2
SF-7 Mental helse (0-4)	0	1.53 (1.14)	1.00 (4)	16.3	8.2
SF-6 Sosial funksjon (0-4)	0	1.55 (1.23)	1.00(4)	24.5	6.1
SF-2 Fysisk funksjon (0-4)	0	2.02 (0.93)	2.00(4)	6.1	2.0
SF-3 Fysisk rollefunksjon (0-4)	0	2.02 (1.05)	2.00(4)	6.1	6.1
SF-5 Overskudd (0-4)	0	2.51 (1.02)	3.00 (4)	2.0	16.3
SF-1 Generell helse (0-5)	0	2.84 (1.23)	3.00 (4)	2.0	6.1
SF-4 Kropps smerte (0-5)	0	3.27 (0.97)	3.00 (4)	0	6.1

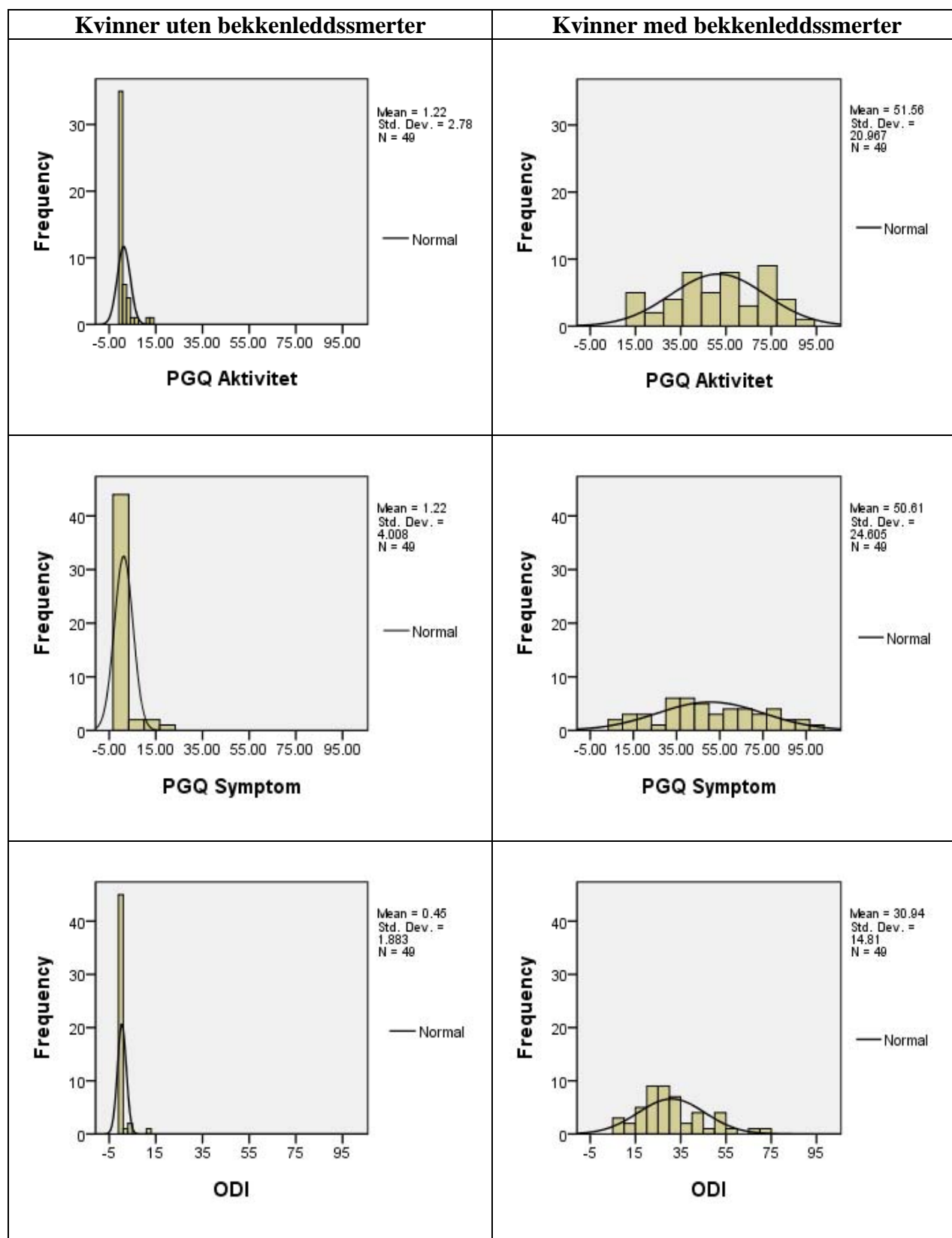
Tabell 4: Skåringer på hvert enkelt utsagn i de tre instrumentene -PGQ-ODI og SF-8- for kvinner uten bekkenleddssmerter; missing data, gjennomsnitt skår med standard avvik (SD), median med interkvartil range(IQR), andeler med laveste og høyeste skår

n= 49	Missing data n(%)	Mean (SD)	Median (IQR)	Lavest skår %	Høvest skår %
PGO aktivitet (0-3)					
1. Kle på deg selv	0	0.00 (0.00)	0.00 (0)	100.0	0
2. Stå mindre enn 10 minutter	0	0.00 (0.00)	0.00 (0)	100.0	0
4. Bøye deg	0	0.00 (0.00)	0.00 (0)	100.0	0
5. Sitte mindre enn 10 minutter	0	0.00 (0.00)	0.00 (0)	100.0	0
7. Gå mindre enn 10 minutter	0	0.00 (0.00)	0.00 (0)	100.0	0
8. Gå mer enn 60 minutter	0	0.00 (0.00)	0.00 (0)	100.0	0
10. Husarbeide	0	0.00 (0.00)	0.00 (0)	100.0	0
11. Bære lett	0	0.00 (0.00)	0.00 (0)	100.0	0
19. Ha et normalt seksualliv	0	0.00 (0.00)	0.00 (0)	100.0	0
9. Gå trapper	0	0.02 (0.14)	0.00 (1)	98.0	0
12. Løfte tungt	0	0.02 (0.14)	0.00 (1)	98.0	0
14. Skyve en vogn	0	0.02 (0.14)	0.00 (1)	98.0	0
17. Ligge	0	0.02 (0.14)	0.00 (0)	98.0	0
18. Snu deg i sengen	0	0.02 (0.14)	0.00 (1)	98.0	0
6. Sitte mer enn 60 minutter	0	0.04 (0.20)	0.00 (1)	100.0	0
13. Reise/sette seg	0	0.04 (0.20)	0.00 (1)	95.9	0
20. Skyve noe med den ene foten	0	0.10 (0.37)	0.00 (2)	91.8	0
15. Løpe	0	0.14 (0.35)	0.00 (1)	85.7	0
16. Utføre sportslige aktiviteter	0	0.14 (0.35)	0.00 (1)	85.7	0
3. Stå mer enn 60 minutter	0	0.16 (0.43)	0.00 (2)	85.7	0
PGO symptom (0-3)					
23. Svikter benet under deg?	0	0.02 (0.14)	0.00 (1)	98.0	0
24. Gjør du ting langsommere?	0	0.02 (0.14)	0.00 (1)	98.0	0
25. Forstyrres nattesøvn?	0	0.02 (0.14)	0.00 (1)	98.0	0
21. Smerter om morgenen	0	0.04 (0.20)	0.00 (1)	95.9	0
22. Smerter om kvelden	0	0.04 (0.20)	0.00 (1)	95.9	0
ODI (0-5)					
4. Gå	0	0.00 (0.00)	0.00 (0)	100	0
2. Personlig stell	0	0.00 (0.00)	0.00 (0)	100.0	0
3. Løfte	0	0.02 (0.00)	0.00 (1)	98.0	0
7. Sove	0	0.02 (0.14)	0.00 (1)	98.0	0
8. Seksualliv	0	0.02 (0.14)	0.00 (1)	98.0	0
6. Stå	0	0.04 (0.20)	0.00 (1)	95.5	0
9. Sosialt liv	0	0.02 (0.14)	0.00 (1)	98.0	0
10. Reising	0	0.02 (0.14)	0.00 (1)	98.0	0
1. Smerteintensitet	0	0.04 (0.20)	0.00 (1)	95.9	0
5. Sitte	0	0.04 (0.29)	0.00 (2)	98.0	0
SF8					
SF-3 Fysisk rollefunksjon (0-4)	1	0.06 (0.32)	0.00 (1)	95.9	0
SF-8 Emosjonell rollefunksjon (0-4)	1	0.13 (0.33)	0.00 (1)	87.5	0
SF-3 Fysisk funksjon (0-4)	1	0.17 (0.43)	0.00 (1)	85.4	0
SF-6 Sosial funksjon (0-4)	1	0.23 (0.43)	0.00 (1)	77.1	0
SF-4 Kroppssmerte (0-5)	1	0.35 (0.67)	0.00 (3)	72.9	0
SF-7 Mental helse (0-4)	1	0.42 (0.58)	0.00 (2)	62.5	0
SF-1 Generell helse (0-5)	1	0.79 (0.74)	1.00 (1)	37.5	0
SF-5 Overskudd (0-4)	1	1.21 (0.77)	1.00 (3)	12.5	0

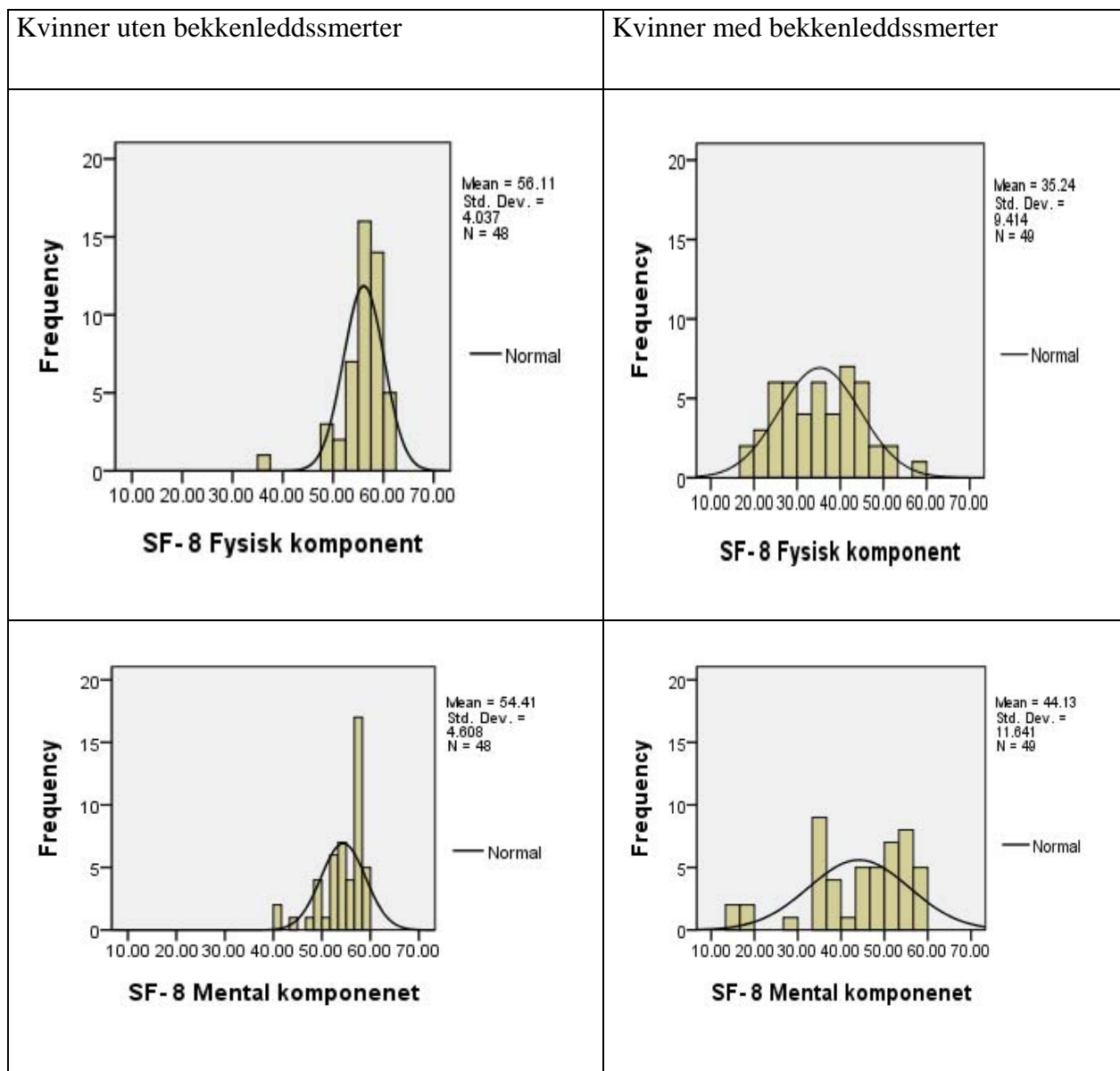
4.3 Fordeling av del - og totalskårer i utfallsmålene

Figur 3 og 4 viser fordeling av henholdsvis del - og totalskårer i de tre instrumentene og for de to gruppene med og uten bekkenleddssmerter. Figurene viser at for alle instrumentene er det en normalfordeling av skårene hos kvinner med bekkenleddssmerter, mens det er en ikke-normal fordeling for kvinnene uten bekkenleddssmerter.

Histogrammene viser at det var ingen gulv- eller takeffekt i noen av instrumentene i gruppen med kvinner med bekkenleddssmerter. I gruppen kvinner uten bekkenleddssmerter viser imidlertid både PGQ og ODI gulv effekt, men det var ikke tilfelle i SF-8 delkomponent skårene.



Figur 3: Histogrammer med normalfordelingskurve av skåringene i PGQ Aktivitet, PGQ Symptom og ODI. Skårene for kvinnene uten bekkenleddssmerter til venstre og gruppen med bekkenleddssmerter til høyre.



Figur 4: Histogrammer med normalfordelingskurve av skåringene SF-8 Fysisk komponent og SF-8 Mental komponent. Skårene for kvinner uten bekkenleddssmerter til venstre og gruppen med bekkenleddssmerter til høyre

4.4 Sammenligning av skårer hos kvinner med og uten bekkenleddssmerter

I figur 5 presenteres resultatene for hvordan kvinner med og uten bekkenleddssmerter skårer på delskalaene av PGQ samt på PGQ total skår. Figuren viser at det er stor spredning rundt median hos kvinnene med bekkenleddssmerter og det betyr at det er stor variasjon i

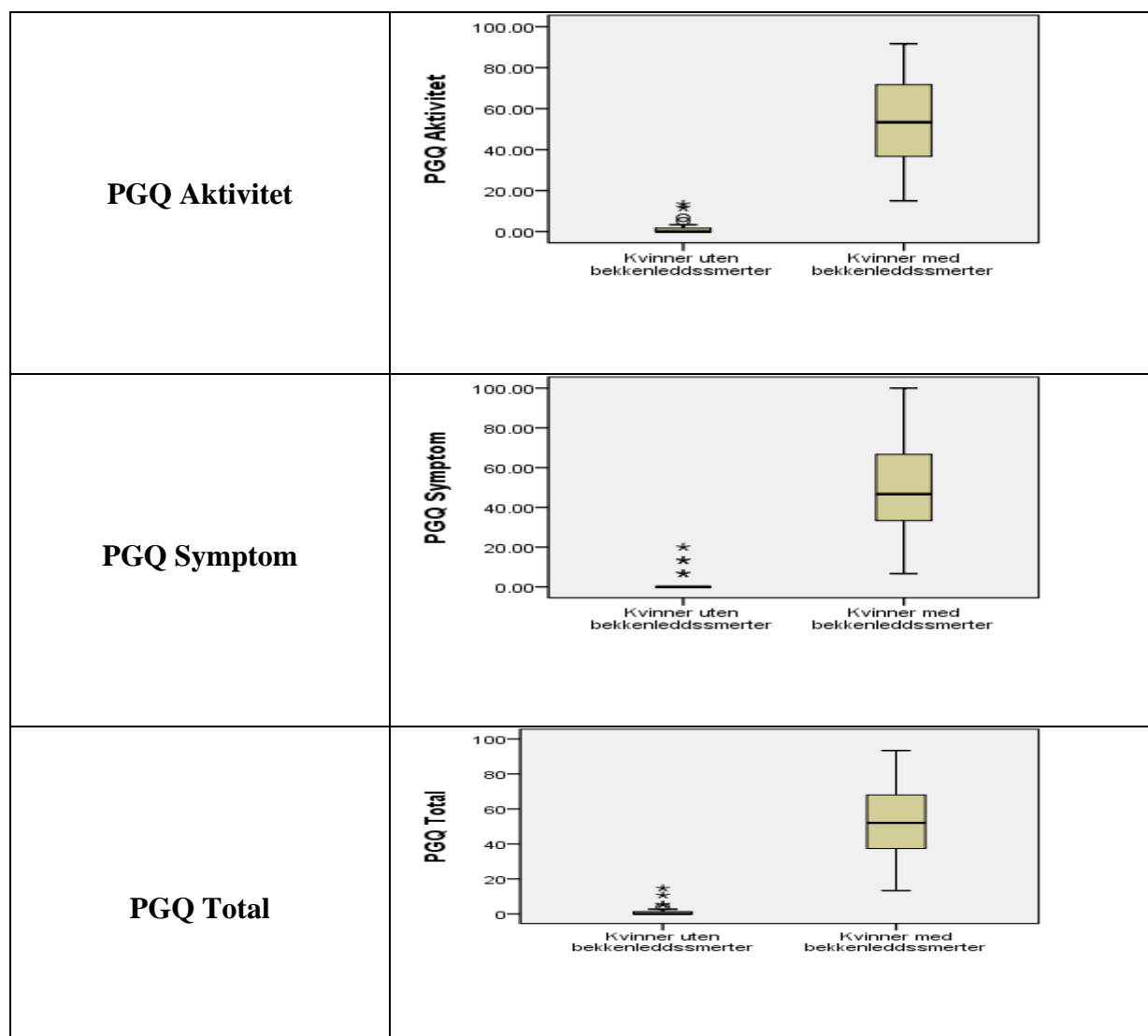
alvorlighetsgraden av bekkenleddssmertene i denne gruppen. Det er liten spredning i skårene til kvinner uten bekkenleddssmerter.

Kvinnene uten bekkeleddssmerter skåret i gjennomsnittlig 1.2 (SD 2.8) på PGQ Aktivitet, mens kvinnene med bekkenleddssmerter skåret 51.6 (SD 21.0). Når det gjelder PGQ Symptom skåret de henholdsvis 1.2 (SD 4.0) og 50.6 (SD 24.6) i gjennomsnitt. Kvinnene uten bekkenleddssmerter skåret i gjennomsnitt 1.12 (SD 2.8) på PGQ Total, mens kvinnene med bekkenleddssmerter skåret 51.37 (SD 21.1).

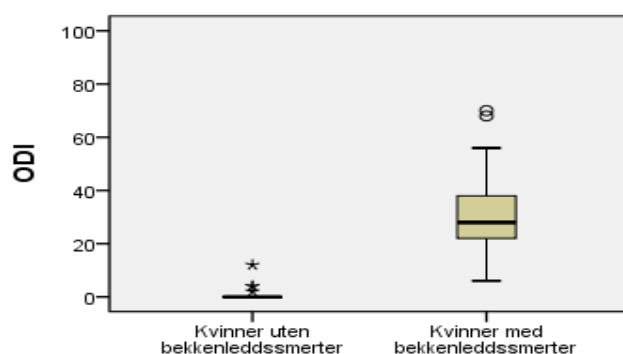
I figur 6 presenteres skåringen av ODI i begge gruppene. Kvinnene uten bekkenleddssmerter skåret i gjennomsnitt 0.45 (SD1.88), mens kvinnene med bekkenleddssmerter scoret i gjennomsnitt 30.94 (SD14.81).

Figur 7 viser Fysisk- og Mental komponent skårene på SF-8 for begge gruppene. Også her skårer kvinnene med bekkenleddssmerter høyere enn kvinnene uten bekkenleddssmerter, men forskjellen er noe mindre enn for PGQ og ODI. Det er større forskjell på den fysiske komponent skåren enn på den mentale.

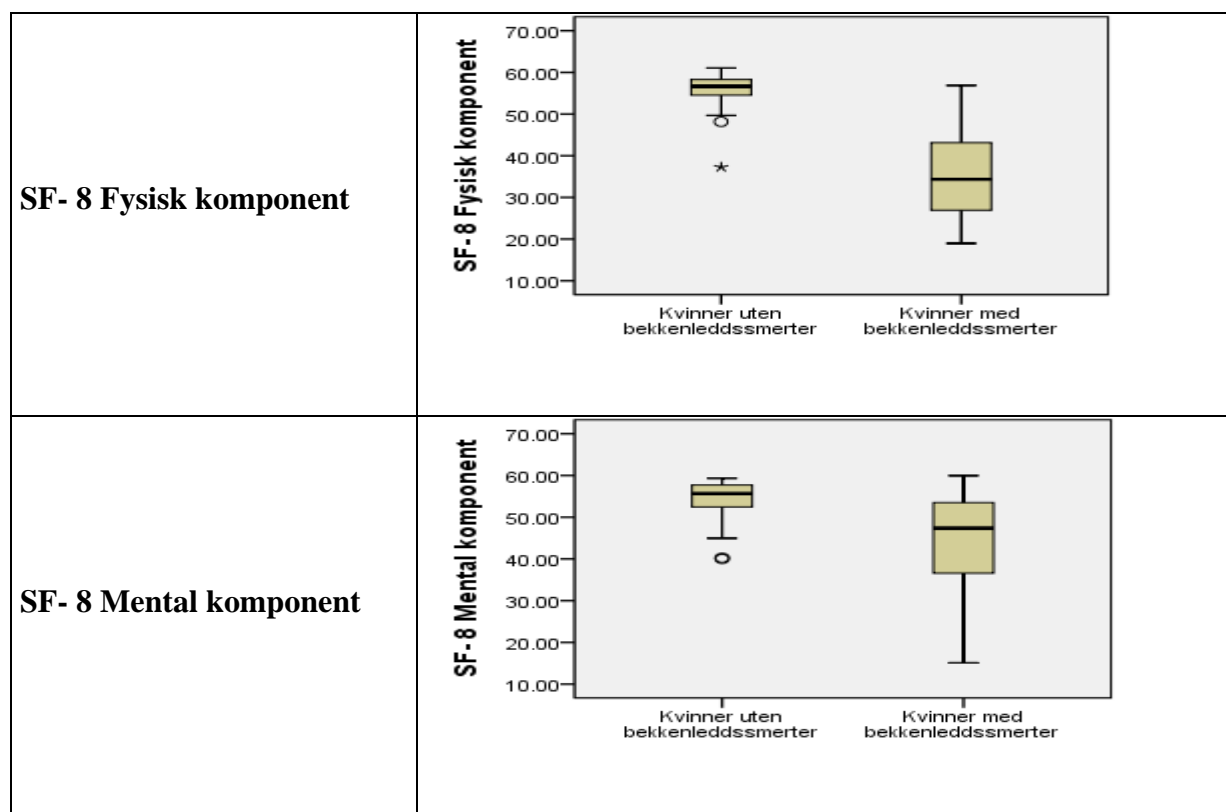
Siden dataene i gruppen med kvinner uten bekkenleddssmerter ikke var normalfordelte ble testingen av forskjell i skårene mellom de to gruppene analysert ved hjelp av Mann-Whitney test, som baserer seg på fordeling av medianskårene. Mann-Whitney testen ga signifikante verdier for alle instrumentene, både del - og totalskårer, med en p verdi $<.001$ (tabell 5).



Figur 5: Median og interkvartil range for PGQ delskalaene og total skåren (0-100) hos kvinner uten bekkenleddssmerter til venstre (n=49) og kvinner med bekkenleddssmerter til høyre (n=49)



Figur 6: Median og interkvartil range for ODI hos kvinner uten bekkenleddssmerter til venstre og kvinner uten bekkenleddssmerter til høyre (n=49)



Figur 7: Median og interkvartil range for SF-8 Fysisk komponent og SF-8 Mental komponent hos kvinner uten bekkenleddssmerter til høyre(n=48) og kvinner med bekkenleddssmerter til venstre(n=49)

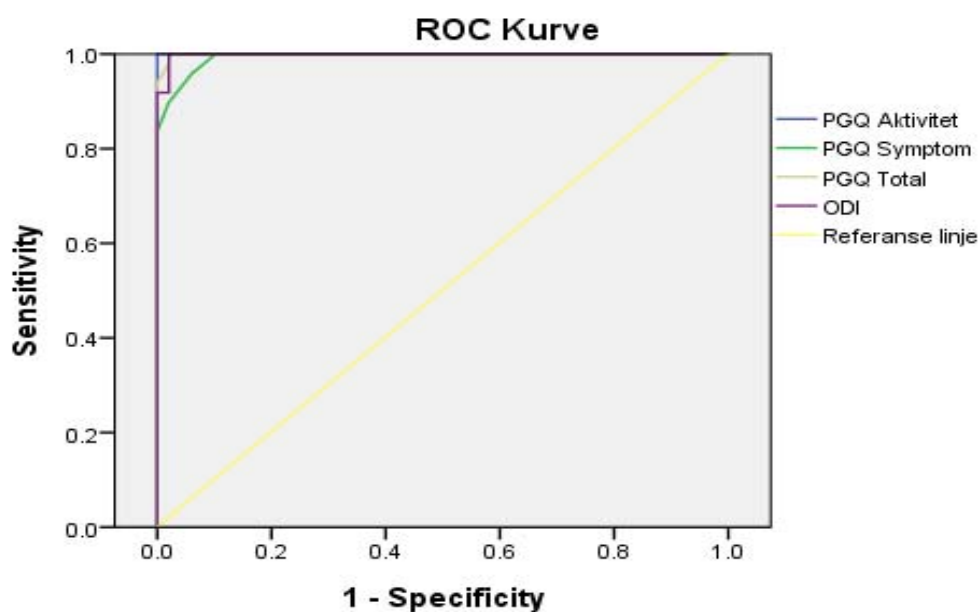
Tabell 5: Median skårer og Mann-Whitney test med interkvartil range (IQR) for PGQ Aktivitet skår, PGQ Symptom skår, PGQ Total skår, ODI, SF-8 Fysisk komponent og SF-8 Mental komponent

	Median (IQR) Kvinner uten bekkenleddssmerter	Median (IQR) Kvinner med bekkenleddssmerter	P verdi
PGQ Aktivitet (0-100)	.00 (13.3)	53.33 (76.7)	<.001
PGQ Symptom (0-100)	.00 (20.0)	46.67 (93.3)	<.001
PGQ Total (0-100)	.00 (15)	52.00 (80)	<.001
ODI (0-100)	.00 (12)	28.00 (64)	<.001
SF-8 Fysisk komponent	55.65 (23.3)	34.34 (37.9)	<.001
SF.8 Mental komponent	55.68 (40.0)	47.36 (44.9)	<.001

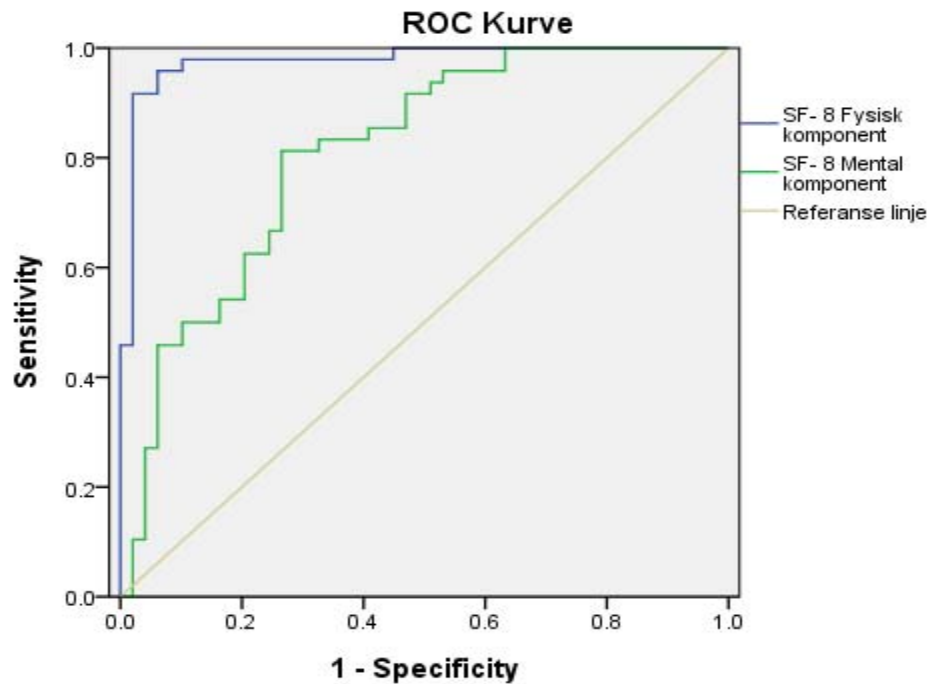
4.5 Diskriminerende evne

Figur 8 og 9 viser ROC kurvene for henholdsvis PGQ, ODI og SF-8 komponent skårene. Jo nærmere kurven er det høyre øvre hjørnet desto bedre skiller instrumentet mellom kvinner med og uten bekkenleddssmerter. I figur 8 er den blå linjen, dvs. PGQ Aktivitet, nærmest det høyre hjørnet, mens i figur 9 er SF-8 Fysiske komponent nærmest.

Tabell 6 viser at arealet under ROC kurven var svært høyt for alle instrumentene. Dette betyr at alle instrumentene skilte svært godt mellom kvinner med og uten bekkenleddssmerter. For eksempel viser tabellen at det er 99 % sjanse for at en kvinne med bekkenleddsplager scorer høyere på PGQ enn en kvinne uten slike plager. Det laveste arealet hadde SF-8 Mental komponent skåren med et areal på .810. De overlappende konfidensintervallene tyder på at det er ingen statistisk signifikant forskjell i diskriminerende evne mellom instrumentene.



Figur 8: Sensitivitet og spesifisitet av Pelvic Girdle Questionnaire (PGQ) Aktivitet, Symptom og Total skår og Oswestry Disability Index (ODI). Jo nærmere kurven går til venstre hjørne, desto bedre sensitivitet og spesifisitet til å skille mellom kvinner uten bekkenleddssmerter og kvinner med bekkenleddssmerter



Figur 9: Sensitivitet og spesifisitet av Short Form – 8 (SF-8) Fysisk komponent og Mental komponent skår. Jo nærmere kurven går til venstre hjørne, desto bedre sensitivitet og spesifisitet til å skille mellom kvinner uten bekkenleddssmerter og kvinner med bekkenleddssmerter

Tabell 6: Areale under ROC kurven, med 95 % konfidensintervall og p-verdi

	Areal under kurven	95% KI	p-verdi
PGQ Aktivitet	1.00	(1.000 – 1.000)	< 0.001
PGQ Symptom	0.994	(0.985 – 1.003)	< 0.001
PGQ Total	0.999	(0.997 – 1.001)	< 0.001
ODI	0.998	(0.994 – 1.002)	< 0.001
SF-8 Fysisk komponent	.977	(.948-1.00)	<.001
SF-8 Mental komponent	.810	(.724-.896)	<.001

5 Diskusjon

I denne masteroppgaven der hovedmålet var å undersøke diskriminerende validitet av tre utfallsmål PGQ, ODI og SF-8 hos kvinner med og uten bekkenleddssmerter etter fødsel viste resultatene at alle utfallsmålene diskriminerte mellom disse to gruppene. Hypotesen om forventede høyere skår hos kvinner med bekkenleddssmerter sammenlignet med kvinner uten bekkenleddssmerter ble dermed støttet av disse resultatene. Studien viste videre en god datakvalitet i form av bortimot ingen manglende data i de to gruppene. Videre viste studien normalfordelte data hos kvinner med bekkenleddssmerter i alle utfallsmålene. Hos kvinnene uten bekkenleddssmerter var det ikke-normalfordelte data og gulveffekt i de to sykdomsspesifikke utfallsmålene PGQ og ODI, mens resultatene viste ikke normalfordelte data og ingen gulveffekt i det generiske instrumentet SF-8. Hypotesene knyttet til delmålene i studien ble dermed støttet av disse resultatene. Studiens hovedresultater vil nå diskuteres i lys av andre relevante studier først, deretter vil jeg diskutere studiens resultater i lys av metodiske styrker og svakheter. I diskusjonens første del inngår også en diskusjon av utfallsmål hos gruppen med bekkenleddssmerter.

5.1 Resultater i lys av andre relevante studier

Denne studien viste at alle utfallsmålene hadde et areale under ROC kurven på mer enn 0.7, som er kriteriet for god diskriminerende evne i henhold til Terwee-gruppen (Terwee et al. 2007). PGQ og ODI, samt den fysiske komponent skåren i SF-8 hadde alle et svært høy areale under kurven på over 0.90. Mental komponent skåren i SF-8 hadde det laveste areale på 0.81.

De overlappende konfidensintervallene tyder imidlertid på at det er ingen statistisk signifikant forskjell i diskriminerende evne mellom de benyttede utfallsmålene i denne studien. De høye og overlappende arealene under ROC kurvene kan sannsynligvis forklares ved at vi sammenlignet to svært ulike grupper, nemlig kvinner med og uten bekkenleddssmerter.

Hadde vi sammenlignet to grupper der begge hadde bekkenleddssmerter ville resultatene sannsynligvis vært mindre overlappende. I studien til Grotle og medarbeidere (Grotle et al. 2012) ble dette gjort, og der var diskriminerende evne mellom ulike subgrupper av bekkenplager mindre enn i vår studie. Grotle et al (Grotle et al. 2012) benyttet også ROC kurver for å undersøke diskriminerende evne mellom gravide og ikke gravide kvinner med bekkenleddssmerter, samt mellom kvinner med ett og flere smertelokalisasjonsområder, og

fant at både PGQ og ODI diskriminerte godt mellom disse subgruppene, mens flere av SF-8 delskalaene manglet diskriminerende evne. Areale under ROC kurvene var mye lavere i Grotles (Grotle et al. 2012) studie enn i denne masteroppgaven, for eksempel varierte areale under ROC kurvene for PGQ delskalaene og ODI fra 0.66 til 0.73, mens for SF-8 var det fra 0.34 opp til 0.63. Gode metodestudier på utfallsmål for kvinner med svangerskapsrelaterte bekkenleddssmerter er mangelfulle. Det er derfor vanskelig å sammenligne denne masteroppgaven med andre studier.

En styrke i denne masteroppgaven er at det ikke ble funnet noen missing, eller manglende data, på noen av skjemaene. Videre var det ikke tak - eller gulveffekt i henhold definisjonen til Terwee (Terwee et al. 2007) på sum skårene av PGQ, ODI eller SF-8 i gruppen med bekkenleddssmerter. Disse resultatene samsvarer med studien til Grotle et al (Grotle et al. 2012), som også fant god datakvalitet på både PGQ, ODI og SF-8. Gjennomsnittsskårene på PGQ og ODI hos kvinnene med bekkenleddssmerter i denne oppgaven var imidlertid gjennomgående noe lavere enn i studiene til Stuge (Stuge et al. 2011) og Grotle (Grotle et al. 2012) med medarbeidere, mens gjennomsnittsskårene for SF-8 var svært like. Forklaringen kan ligge i forskjellen i utvalgene. Stuge og medarbeidere (Stuge et al. 2011) hadde i sin studie med gravide kvinner, mens det i denne masteroppgaven kun hadde kvinner med bekkenleddssmerter etter fødsel. Studien til Stuge og medarbeidere (Stuge et al. 2011) viste at gravide kvinner med bekkenleddssmerter hadde problemer med de samme aktivitetene som kvinnene med bekkenleddssmerter etter fødsel. Imidlertid hadde de gravide kvinnene høyere grad av aktivitetsproblemer sammenlignet med de ikke gravide kvinnene.

Tak - eller gulveffekt av sum skår bør ikke blandes med det faktum at en betydelig del av kvinnene med bekkenleddssmerter skåret på laveste og høyeste mulig skår for mange av utsagnene i utfallsmålene. Lave skår på lette utsagn (for eksempel gå mindre enn 10 minutter) er rimelig å forvente når personen har lite funksjonsproblemer, mens høye skår på lette utsagn vil derimot reflektere en person med store funksjonsproblemer. Det motsatte gjelder lave skår på vanskelige utsagn (for eksempel løpe). Ved å se nærmere på dette skåringsmønsteret kan man få et inntrykk av hvordan utsagnene i de ulike utfallsmålene dekker det fenomenet som det er ment å måle, noe som kan analyseres i en Rasch analyse. Fordelingen av lette og vanskelige utsagn i PGQ delskalaene i denne studien samsvarte godt med resultatene fra Rasch analysene under utviklingen av PGQ (Grotle et al. 2012).

Hvis man sammenligner gjennomsnittsskårene på de to sykdomsspesifikke utfallsmålene i gruppen med bekkenleddssmerter i denne masteroppgaven så skåret disse kvinnene gjennomsnittlig 51 på PGQ Aktivitet og 31 på ODI, begge målt på en 0-100 skala. Dette kan tyde på at PGQ Aktivitet er mer sensitiv til å måle aktivitetsbegrensning for kvinner med bekkenleddssmerter enn ODI. Dette samsvarer med resultatene i studien til Grotle og medarbeidere (2012). Denne forskjellen kommer også til syne i skåringen av utsagnet ”seksualliv”, som er et utsagn i både PGQ og ODI. Mens ingen skårer på den høyeste delen av skalaen på ODI, er det 10 % av kvinnene som skårer på den høyeste delen av skalaen på PGQ. Det er mulig at formuleringen av utsagnene kan påvirke denne forskjellen i skåringsmønstre, eventuelt at kvinnene føler at PGQ er mer rettet mot deres problemer enn ODI.

I gruppen uten bekkenleddssmerter ble det funnet svært lave skåringer og en klar gulveffekt i PGQ og ODI, noe som mest sannsynlig forklares med at disse kvinnene hadde ingen bekkenleddssmerter eller symptomer etter fødsel og heller ingen aktivitetsbegrensninger. På det generiske skjemaet SF-8 hvor høyt skår betyr bra helse skåret kvinnene uten bekkenleddssmerter høyt. Dette funnet er i samsvar med en annen norsk studie der SF-36 ble benyttet i en normalpopulasjon og hvor det ble funnet en tak-effekt på flere av subskalaene (Loge and Kaasa 1998). SF-8 inneholder som tidligere nevnt de samme subskalaene som SF-36, og kan derfor sammenliknes med denne studien.

5.2 Utfallsmål for bekkenleddssmerter

5.2.1 PGQ

Så langt jeg vet, er PGQ det første sykdomsspesifikke funksjonsskjema som er spesielt utviklet for kvinner med bekkenleddssmerter og det er dokumentasjon for at PGQ er egnet for denne pasientgruppen (Stuge et al. 2011). En styrke ved dette skjemaet er at kvinner med verifiserte bekkenleddssmerter har bidratt i utviklingen av skjemaet. Studier har vist at det er forskjell i klinikerens og pasienters oppfatning om virkning av sykdom, behandlings utfallets prioriteringer og verdier av ulike utfall (Hewlett et al. 2001; Kvien and Heiberg 2003).

Tradisjonelt er kliniske tiltak ofte basert på klinikerens meninger og mangler derfor innholdsvaliditet fra pasient perspektivet. Pasient perspektivet om hvordan deres tilstand eller sykdom påvirker livene deres må vurderes når man utvikler et effektmål som er designet for å

vurdere helsetilstanden og resultater av betydning for mottakere av helsetjenester. Innholdet i PGQ ble basert på informasjon fra fokus gruppe og innspill fra erfarende fysioterapeuter som daglig jobber med kvinner med bekkenleddssmerter. Alvorlighetsgraden av symptomer er en annen faktor som kan påvirke relevansen av ulike utfall (Casarett et al. 2001). Under utviklingen av PGQ ble det sørget for at utvalget i fokusgruppen hadde et bredt spekter av alvorlighetsgrad av bekkenleddssmerter. Utvalget bestod av både gravide og kvinner etter fødsel, alle med bekkenleddssmerter. Pasientgruppen var et representativt utvalg av kvinner med bekkenleddssmerter hva gjelder alder og varighet av bekkenleddssmerter. Samarbeidet mellom fysioterapeuter som jobber med pasienter med bekkenleddssmerter og forskere som har erfaring med helsestatus skjemaer var med på å sikre at klinikken og forskningen ble ivarettatt på en god måte. I tillegg var svarprosenten veldig høy. Kvinner med bekkenleddssmerter har også i tidligere studier vist å ha høy svarprosent (Stuge et al. 2004a).

Tidligere studier av bekkenleddssmerter har brukt helsestatus skjemaer som er utviklet for pasienter med rygg smerter og de er kanskje ikke adekvate for pasienter med bekkenleddssmerter (Bastiaenen et al. 2006; Nilsson-Wikmar et al. 2005; Stuge et al. 2004a). Bekkenleddssmerter er sett på som en tilstand som er annerledes enn rygg smerter (Vleeming et al. 2008). Det er derfor mulig at funksjons nedsettelse hos pasienter med bekkenleddssmerter vil være annerledes enn for pasienter med rygg smerter. Man fant imidlertid i studien til Stuge og medarbeidere (2011) at mange utsagn var overlappende. Utsagn som ligner på dem som er brukt i PGQ er også funnet i funksjonsskjemaene som er utviklet for rygg pasienter, men alle utsagnene som er benyttet i PGQ er ikke funnet i ett skjema.

Under pilottestingen av PGQ ble det identifisert 3 aktiviteter som ikke er inkludert i utfallsmålene utviklet for ryggpasienter; skyve en vogn, svikter benet under deg og skyve noe med den ene foten. "Å skyve noe med den ene foten" ligner "Pulling a mat" test som er beskrevet i en svensk studie (Fagevik et al. 2009). "Pulling a mat" er en test hvor pasienten utfører hofte abduksjon og adduksjon. Denne testen er funnet å være meget sensitiv for fremre bekkenleddssmerter (Fagevik et al. 2009). "Å skyve noe med den ene foten" vil sannsynligvis fange opp de pasientene som har plager relatert til symfyse. The Patient-Specific Functional Scale som ble brukt i utviklingen av PGQ (Stuge et al. 2011) viste at 61 % av deltagerne (73 % gravide og 48 % ikke gravide) i studien rapporterte problemer med å gå. Lignende er funnet i en annen studie hvor 81 % av pasientene med bekkenleddssmerter rapporterte smerter

med gange (Mens et al. 1996). En vesentlig forskjell mellom PGQ og ODI er at utsagnene gå, stå og sitter er delt inn i to kategorier i PGQ; ”stå mindre enn 10 minutter” og ”stå mer enn 60 minutter”. Dette gir rom for en bedre gradering av disse aktivitetene som igjen kan gjenspeile skåringene på de to skjemaene.

Fysioterapeuter behandler ofte pasienter med bekkenleddssmerter under graviditeten og etter fødsel. Derfor er det behov for helsestatus skjema som kan brukes for begge tilstandene (Boissonnault 2009). Studien til Stuge og medarbeidere (2011) viste at deltagerne hadde problemer med de samme aktivitetene både under graviditeten og etter fødsel. Graden av vanskelighet var imidlertid forskjellige hos dem som var gravide og de som ikke var gravide (Stuge et al. 2011). PGQ er det eneste helsestatus skjemaet som viste tilfredsstillende diskriminerende validitet mellom gravide med bekkenleddssmerter og de som ikke var gravide med bekkenleddssmerter (Grotle et al. 2012), det vil si at det er sensitivt for alvorlighetsgrad av bekkenleddssmerter.

Kvinnene med bekkenleddssmerter skåret i gjennomsnitt vesentlig mye høyere på PGQ sammenlignet med kvinnene uten bekkenleddssmerter. PGQ er som tidligere nevnt et nylig utviklet funksjonsskjema og det er derfor få studier å sammenligne med. Det eneste man kan sammenligne med er materialet skjemaet ble utviklet på (Grotle et al. 2012; Stuge et al. 2011). Det finnes ikke normal data å sammenligne med da PGQ ikke er testet ut på normal populasjon tidligere.

5.2.2 ODI

For å teste diskriminerende evne av PGQ var det viktig å kunne sammenlikne det med andre hyppig brukte utfallsmål. I en oversiktsartikkel fant Grotle og medarbeidere (Grotle et al. 2004b) 36 forskjellige utfallsmål. Roland & Morris Disability Questionnaire (RMDQ) (Roland and Fairbank 2000) og ODI var de mest brukte og anbefalte utfallsmålene. RMDQ har vist seg å være mer sensitiv til endring i pasienter med lette til moderate plager, ODI ser ut til å være mer sensitiv for pasienter med alvorlige plager (Grotle et al. 2004a). Kvinner med vedvarende bekkenleddssmerter kan i noen tilfeller ha plager av alvorlig grad (Stuge et al. 2004a). Det kan være en av forklaringene til at ODI har vært det skjemaet som har vært hyppigst benyttet i kliniske studier av kvinner med bekkenleddssmerter (Gutke et al. 2006; Stuge et al. 2004a).

Det var stor forskjell på skåringen på ODI i gruppen med bekkenleddssmerter og de uten bekkenleddssmerter. Kvinnene med bekkenleddssmerter skåret i gjennomsnitt 31 poeng som tilsvarer moderate funksjonsproblemer. Dette var noe lavere sammenlignet med materiale til Stuge og medarbeidere (2011). Noe av denne forskjellen kan forklares på bakgrunn av at materialet til Stuge (Stuge et al. 2011) også hadde med gravide kvinner med bekkenleddssmerter. I en randomisert kontrollert studie av kvinner med klinisk verifiserte bekkenleddssmerter etter fødsel (Stuge et al. 2004a) skåret kvinnene i gjennomsnitt 42 poeng på ODI, noe som tilsvarer betydelige funksjonsproblemer. I en studie på svenske gravide i 12.-18. uke, skåret gruppen med bekkenleddssmerter 14 poeng på ODI, mens kvinner med både bekkenleddssmerter og ryggsmarter skåret 18 poeng (Gutke et al. 2006). Kvinnene i Gutke's studie hadde minimale funksjonsproblemer. Gutke's materiale ble rekruttert via helsestasjoner hvor gravide kvinner ble bedt om å fylle ut skjema mens de var på svangerskaps kontroll. Disse kvinnene hadde ikke oppsøkt behandling for sine plager. I studien av Stuge og medarbeidere (2004a) derimot hadde man et materiale på kvinner som søkte behandling for bekkenleddssmerter etter fødsel. Stuges materiale var en gruppe pasienter som ikke var blitt bra etter fødsel og representerte antagelig en dårlig gruppe. Flere studier viser at de aller fleste blir helt bra av sine bekkenleddsplager de første månedene etter fødsel (Bjelland et al. 2012; Ostgaard et al. 1996; Robinson et al. 2010b). Rekrutteringen av pasientene til de forskjellige studiene kan påvirke alvorlighetsgraden av plagene hos kvinnene med bekkenleddssmerter.

Kvinnene uten bekkenleddssmerter i denne masteroppgaven skåret i gjennomsnitt 0.45 poeng som indikerer at de ikke hadde noen grad av funksjonsnedsettelse. Dette er ikke i samsvar med Fairbank og Pynsent (2000). I en oversiktsartikkel fant Fairbank og Pynsent (2000) en normal skår på 10.19 for ODI. Denne konklusjonen er bla. basert på to små biomekaniske studier som brukte en kontroll gruppe som ikke hadde smerter i ryggen (Kankaanpaa et al. 1997; Kankaanpaa et al. 1998). En gruppe i Japan estimerte normal skåring på ODI til å være 8.7 og en cut - off skår på 12 (Tonosu et al. 2012). Gutke og medarbeidere (2006) hadde imidlertid i sin studie 64 gravide i uke 12-18 uten bekkenleddssmerter. Denne gruppen hadde en gjennomsnitt skår på 4 poeng. Studien til Gutke (2006) er mer i samsvar med vår studie. Det er naturlig at "friske" gravide skårer litt høyere enn "friske" ikke gravide.

5.2.3 SF- 8

SF-8 er så vidt jeg vet ikke benyttet i tidligere studier av kvinner med klinisk verifiserte bekkenleddssmerter. Eggen og medarbeidere (2012) benyttet imidlertid SF-8 i en studie med kvinner med selvrapporterte bekkenplager. De rapporterte Fysisk komponent skår og Mental komponent skår på en gruppe gravide kvinner. Materialet i Eggen`s studie skåret høyere på den Fysiske komponent skåren sammenlignet med kvinnene med bekkenleddssmerter i vår studie, mens skåringene var mer i samsvar på den Mentale komponent skåren. Dette kan tyde på at gravide uten bekkenleddssmerter har bedre funksjon enn ikke gravide kvinner med bekkenleddssmerter. Kvinnene uten bekkenleddssmerter i vår studie skåret veldig høyt på alle delskårene på SF-8. Dette er i samsvar med studien til Loge (Loge and Kaasa 1998) hvor de testet ut SF-36 på normal befolkning i Norge. SF-8 kan som tidligere nevnt sammenlignes med SF-36 skårene for de ulike subskalaene (Lefante, Jr. et al. 2005). Loge fant imidlertid også at utdannings nivå påvirket alle subskalaene, forskjellen var minst på de mentale subskalene (Loge and Kaasa 1998). Utdanningsnivået kan være en medvirkende årsak til at kvinnene uten bekkenleddssmerter i vår studie skåret veldig høyt på Funksjons komponent skåren. Forskjellen mellom kvinnene med bekkenleddssmerter og uten bekkenleddssmerter i vår studie var imidlertid størst på den Fysiske komponent skåren sammenlignet med den Mentale komponent skåren.

I utviklingen av PGQ (Grotle et al. 2012) benyttet man også SF-8. Dette materialet består av både gravide med bekkenleddssmerter og ikke gravide med bekkenleddssmerter.

Skåringsprofilen på materialet til Grotle og medarbeidere (2012) ligger stort sett litt lavere enn gruppen med bekkenleddssmerter i vår studie. Unntaket er "overskudd" og "emosjonell rollefunksjon" hvor materialet vårt ligger litt over. "Overskudd" er det domenet som skiller seg ut i materialet til Grotle (2012) og i begge gruppene i denne studien. Det er stor andel av småbarns mødre i begge materialene. Det å være småbarns mødre er krevende enten du er frisk eller syk.

Det står på hjemmesiden til SF-8 at SF-8 ikke egner seg for små populasjoner (SF8(tm)Health Survey). Imidlertid finner jeg i denne masteroppgaven at SF-8 skiller godt mellom kvinner med bekkenleddssmerter og kvinner uten bekkenleddssmerter i vårt lille materiale. Det er dermed grunn til å anta at det likevel kan benyttes i mindre populasjoner.

5.3 Resultater i lys av metodiske aspekter

Det er styrker og svakheter denne masteroppgaven, som kan ha påvirket resultatene. Utvalget av kvinner med bekkenleddssmerter er representativt sammenlignet med andre studier på denne pasientgruppen (Grotle et al. 2012; Stuge et al. 2004a; Stuge et al. 2011). Det er også en styrke at man lyktes med å rekruttere kvinner uten bekkenleddssmerter som matchet kvinnene med bekkenleddssmerter. Videre er det en styrke at alle kvinnene med bekkenleddssmerter ble diagnostisert til å ha bekkenleddssmerter med kliniske tester.

Kvinnene med bekkenleddssmerter ble i hovedsak rekruttert via behandlere på institutter og gruppetreninger for kvinner med bekkenleddssmerter. Det ble også rekruttert fra landsforeningen for kvinner med bekkenløsningssmerter i Oslo. I tillegg var det noen få som tok kontakt på eget initiativ på grunn av oppslag på helsestasjoner og fysikalske institutter. De fleste ble rekruttert gjennom helsepersonell som hadde vurdert pasienten til å ha bekkenleddssmerter. Det var bare noen få som ble rekruttert basert på egenrapporterte plager. Ved å rekruttere på denne måten fikk man tak i kvinner som hadde tatt initiativ til å oppsøke hjelp for plagene sine. Åtti-åtte prosent av kvinnene hadde mer enn 12 års utdanning og det er sammenlignbart med andre norske studier utført på denne pasientgruppen (Robinson et al. 2010a; Stuge et al. 2004a). Det høye utdanningsnivået indikerer at materialet representerer en gruppe ressurssterke kvinner. Kvinnene uten bekkenleddssmerter ble matchet med kvinnene med bekkenleddssmerter hva gjelder alder og antall vaginale fødsler. Nitti-fire prosent av kvinnene i kontrollgruppen hadde mer enn 12 års utdanning. Hovedsakelig ble de rekruttert gjennom venner og kjente av de inkluderte kvinnene med bekkenleddssmerter. Det er derfor god grunn til å anta at disse kvinnene er like ressurssterke og matcher gruppen med bekkenleddssmerter også på dette området.

I denne oppgaven ble det vektlagt å rekruttere kvinner som hadde debut av bekkenleddssmerter under svangerskapet eller 3 uker etter fødsel og minimum 6 måneders varighet. Ved å ekskludere de som hadde mindre enn 6 måneders varighet av bekkenleddssmerter sikret man å få de som var alvorlig plaget. Krav om gode norsk kunnskaper innebærer at man kun fikk en gruppe med kvinner som er godt integrert i det norske samfunn. Dette resulterte i at jeg kanskje ikke fikk en gruppen med kvinner som er representativt for hele befolkningen i Oslo. Det var imidlertid nødvendig at kvinnene hadde gode norsk kunnskaper for å sikre god kommunikasjon og dermed mest mulig pålitelige svar på undersøkelsen.

Selvrappporterte plager er ikke godt nok diagnostisk til å skille mellom rygg- og bekken. For å være helt sikre på at kvinnene med bekkenleddssmerter virkelig var plaget med smerter i bekkenet, måtte man gjøre noen valg i forhold til diagnostiseringen. Det finnes i dag ingen gullstandard for diagnostisering av bekkenleddssmerter. I tidligere studier hvor undersøkelsesprosedyrer for å diagnostisere bekkenleddssmerter er blitt beskrevet er det ofte en kombinasjon av metoder; inspeksjon av gange, holdning og bekken stilling, palpasjon av ligamenter og muskler, test for låst iliosacral ledd og smerteprovokasjonstester for iliosacral leddet og symfysen (Vleeming et al. 2008). Det har i tidligere studier blitt fokusert mest på inspeksjon og palpatoriske funn, mens de senere studiene har fokusert mer på smerteprovokasjonstester, antageligvis fordi de har høyere reliabilitet og spesifisitet (Vleeming et al. 2008). De fleste av de evaluerte testene, og alle testene Europeiske Guidelines anbefaler har høy spesifisitet som betyr at hvis de er negative så er det høy sannsynlighet for at pasientene ikke lider av bekkenleddssmerter. Sensitiviteten på testene er lavere og derfor er det anbefalt å bruke flere tester. Da det ikke finnes noen gullstandard på å diagnostisere bekkenleddssmerter er det vanskelig å vurdere validiteten (Vleeming et al. 2008). I denne masteroppgaven valgte man å følge retningslinjene fra Europeiske Guidelines (Vleeming et al. 2008) som anbefaler at man klinisk tester kvinnene med bekkenleddssmerter for å være helt sikre på at disse kvinnene hadde plager i bekkenet.

Positiv P4 test og/eller ASLR test (skår over 3) og minst en positiv av de andre testene ble benyttet for å øke sannsynligheten for at alle kvinnene var affisert av bekkenleddssmerter. ASLR cut - off på 3 har blitt brukt i andre studier på denne pasientgruppen (Stuge et al. 2004a; Vollestad and Stuge 2009). Mens og medarbeidere (Mens et al. 2001) fant imidlertid at den beste balansen mellom spesifisitet og sensitivitet var når en skår på 0 var negativ og 1-10 ble regnet som positiv. De konkluderte likevel med at en høyere cut - off i noen tilfeller er å foretrekke. Ved cut-off på ASLR mellom 2 og 3 er sensitiviteten 0.66 og spesifisiteten er 100 % (Mens et al. 2001). I vår studie har vi valgt en cut-off på 3 som da gir en spesifisitet på 100 % som betyr at det er liten sjanse for falske positive. Vi var ute etter å være helt sikre på at de kvinnene som ble inkludert til gruppen med bekkenleddssmerter virkelig hadde bekkenleddssmerter og ikke bare ubehag. Mens og medarbeidere (Mens et al. 2012) har i en nylig publisert studie konkludert med at den beste cut-off for ASLR er 1, men det kan stilles spørsmål om hvordan han har diagnostisert pasientene inn i grupper med og uten kombinasjon av rygg smerter i den studien. Forøvrig er antallet lite i de forskjellige undergruppene (Mens

et al. 2012). Ved høyre cut – off som er valgt i vår studie er man sikrere på at man får tak i de som virkelig har bekkenleddssmerter og av en mer alvorlig grad.

I kliniske undersøkelser vil det alltid være elementer av subjektiv vurdering både fra pasient og undersøker. Det ble i vår studie lagt vekt på å standardisere informasjon og praktisk utførelse av testene for å være sikker på at pasientene ble diagnostisert best mulig.

Fysioterapeut Kaja Sætre testet alle pasientene. Hun ble lært opp av fysioterapeut og veileder Britt Stuge i de kliniske testene så vel som muntlig informasjon. En styrke ved denne masteroppgaven er at man hadde en person som utførte alle testene. Dette sikret at man fikk mest mulig lik gjennomføring av testene og registreringen.

De fleste bekkentestene skåres på grunnlag av ja- eller nei svar på spørsmål om smerte. Dette er en subjektiv vurdering som kan gi redusert pålitelighet av testene. I en annen studie har smerteprovokasjonstester blitt skåret med gradering av smerte i flere kategorier (Vleeming et al. 2002). Dette vil kunne gi en mulighet til å få frem en mer nyansert registrering for disse testene, men denne graderingen er ikke reliabilitetstestet og ikke benyttet i øvrige studier på bekkenleddssmerter.

I klinisk arbeid og i forskninger er det en utfordring å differensiere mellom tilstander som har sin årsak i lumbalcolumna og de som primært har bekkenleddssmerter. Nerverotkompresjon er betraktet som et uttrykk for alvorligere grad av lumbal patologi. Nervestrekktester vil som ofte gi utstrålende smerter ved slike tilstander (Kerr et al. 2002) (Laslett et al. 2005b). Mange har benyttet denne type tester som sentrale eksklusjonskriterier i vitenskapelige studier relatert til bekkenleddssmerter (Albert et al. 2000; Laslett et al. 2005b; Mens et al. 2000). Nervestrekktester alene er likevel ikke nok til å skille ut pasienter med lumbal patologi som primær årsak til smertene. Derfor ble nervestrekktesten benyttet sammen med smertetegning som eksklusjon/inklusionskriterier i denne studien. Smertetegning eller intervju for å lokalisere smerte er benyttet i tidligere studier på denne pasientgruppen (Laslett et al. 2003; Laslett et al. 2005b; Mens et al. 2001; Ostgaard et al. 1994; Robinson et al. 2010a; Stuge et al. 2004a).

Femti-seks pasienter ble testet og 51 ble inkludert i studien. Det var 3 pasienter som ble ekskludert på grunn av mangelfulle positive tester og to på grunn av tegn til nevrologi. I tillegg var det 2 som ikke dukket opp til ultralyd testingen og dermed droppet ut. Dette kan

være en indikasjon på at eksklusjons- og inklusjonskriteriene var gode nok til å fange opp pasienter med ryggplager og andre som ikke hadde smerter relatert til bekkenleddene.

En annen styrke med denne masteroppgaven er at jeg har brukt en anbefalt og god metode for å vurdere diskriminerende validitet. ROC analyser er den optimale testen for å analysere i hvilken grad et instrument diskriminerer mellom to tilstander, her mellom friske og ikke-friske kvinner etter fødsel. I henhold til Terwee (Terwee et al. 2007) og COSMIN gruppen er et ROC areale under kurven på 0.70 eller mer et tilfredsstillende kriterium for god diskriminerende evne. Som tidligere nevnt, oppnådde alle utfallsmålene i denne oppgaven dette kriteriet. Videre ble det brukt ikke-parametrisk t-test siden dataene ikke var normalfordelte i gruppen uten bekkenleddssmerter. Dette må også ses på som en styrke ved oppgaven.

En svakhet med oppgaven kan være at vi benyttet to sykdomsspesifikke utfallsmål der formuleringene var relatert til smerte. Således ble kvinnene uten bekkenleddssmerter oppfordret til å svare på to skjemaer som ikke var aktuelle eller relevante for dem. Dette ble diskutert mye før oppstart av studien, og vi landet på dette designet fordi det var den eneste muligheten til å få normalmateriale for det nye instrumentet PGQ. Gitt denne situasjonen var det svært viktig å få med et generisk instrument. Man kunne argumentere med at hele SF-36 burde være benyttet. Det ville gitt et bedre grunnlag for å sammenlikne skåringene hos kvinnene uten bekkenleddssmerter med andre normalmaterialer i Norge.

Som tidligere nevnt er det foreslått et "core set" med utfallsmål for pasienter med rygg plager. Det foreslåtte "core settet" dekker følgende fem domener; spesifikke funksjonsskjema, generiske helsestatus skjemaer, smerte, arbeidsstatus, og pasient tilfredshet. I denne masteroppgaven dekker jeg generiske og sykdomsspesifikke utfallsmål, samt arbeidsstatus, men mangler måling av smerte og tilfredshet i de to gruppene. En måling av smerte kunne ha gitt mer nyansert informasjon om hvor smerteplaget kvinnene med bekkenleddssmerter var, samt man kunne sammenligne med andre smertepopulasjoner slik som rygg smerter. Måling av arbeidsstatus og tilfredshet kan imidlertid være viktig i fremtidige kliniske studier der man undersøker effekt av intervensjoner.

6 Konklusjon

Hensikten med denne studien var å undersøke diskriminerende validitet av de tre utfallsmålene; PGQ, ODI og SF-8 hos kvinner med og uten bekkenleddssmerter. I tillegg ble datakvalitet, samt eventuell tak - eller gulveffekt i skåringene, undersøkt.

Resultatene viste god diskriminerende validitet av alle tre utfallsmålene hos kvinner med bekkenleddssmerter og kvinner uten bekkenleddssmerter. Videre viste studien normalfordelte data og ingen tak- eller gulveffekt hos kvinnene med bekkenleddssmerter i alle utfallsmålene. Hos kvinnene uten bekkenleddssmerter var det ikke normalfordelte data på alle tre utfallsmålene. PGQ og ODI viste gulveffekt hos denne gruppen. Det var imidlertid ikke tilfelle i fysisk komponent og mental komponent skårene for SF-8.

PGQ, ODI og SF-8 er utfallsmål som har vist seg og være reliable og valide når de er testet ut på kvinner med bekkenleddssmerter. PGQ er utviklet for kvinner med bekkenleddssmerter som har oppstått under graviditet eller etter fødsel og det er derfor grunn til å tro at det egner seg godt for denne gruppen. Videre forskning av responsiveness for endring av PGQ er imidlertid nødvendig for å kunne komme med endelig anbefaling av det best egnede utfallsmål for kvinner med bekkenleddssmerter.

Litteraturliste

Albert H, Godskesen M & Westergaard J. Evaluation of clinical tests used in classification procedures in pregnancy-related pelvic joint pain. *Eur.Spine J.* 2000; 9(2): 161-166

Albert H, Godskesen M & Westergaard J. Prognosis in four syndromes of pregnancy-related pelvic pain. *Acta Obstet.Gynecol.Scand.* 2001; 80(6): 505-510

Albert HB, Godskesen M & Westergaard JG. Incidence of four syndromes of pregnancy-related pelvic joint pain. *Spine* 2002; 27(24): 2831-2834

Aldabe D, Milosavljevic S & Bussey MD. Is pregnancy related pelvic girdle pain associated with altered kinematic, kinetic and motor control of the pelvis? A systematic review. *Eur.Spine J.* 2012a; 21(9): 1777-1787

Aldabe D, Ribeiro DC, Milosavljevic S & Dawn BM. Pregnancy-related pelvic girdle pain and its relationship with relaxin levels during pregnancy: a systematic review. *Eur.Spine J.* 2012b; 21(9): 1769-1776

Altman, DG *Practical Statistics for Medical Research.* Chapman & Hall; 1991. p. 1-611.

BAKER, D. J., PYNSET, P. B. & FAIRBANK, J. C. T. (1990) The Oswestry Disability Index revisited: it's reliability, repeatability and validity, and a comparison with the St Thomas' s Disability Index. *In: Roland MO, Jenner JR, eds. Back pain. New approaches to rehabilitation and education.*, Manchester University Press.

Bastiaenen CH, de Bie RA, Wolters PM et al. Effectiveness of a tailor-made intervention for pregnancy-related pelvic girdle and/or low back pain after delivery: short-term results of a randomized clinical trial [ISRCTN08477490]. *BMC.Musculoskelet.Disord.* 2006; 7(19): 1-13

Batterham AM & George KP. Reliability in evidence-based clinical practice: A primer for allied health professionals. *Physical Therapy in Sport* 2003; 4 122-128

Beales DJ, O'Sullivan PB & Briffa NK. Motor control patterns during an active straight leg raise in chronic pelvic girdle pain subjects. *Spine* 2009a; 34(9): 861-870

Beales DJ, O'Sullivan PB & Briffa NK. Motor control patterns during an active straight leg raise in pain-free subjects. *Spine (Phila Pa 1976.)* 2009b; 34(1): E1-E8

Beurskens AJ, de Vet HC & Koke AJ. Responsiveness of functional status in low back pain: a comparison of different instruments. *Pain* 1996; 65(1): 71-76

- Bjelland E, Eberhard-Gran M, Nielsen C & Eskild A. Age at menarche and pelvic girdle syndrome in pregnancy: a population study of 74 973 women. *BJOG*. 2011;
- Bjelland E, Stuge B, Engdahl B & Eberhard-Gran M. The effect of emotional distress on persistent pelvic girdle pain after delivery: a longitudinal population study. *BJOG*. 2012;
- Boissonnault JS. A review of self-report functional outcome measures in selected obstetric physical therapy interventions. *Journal of Women's Health Physical Therapy* 2009; 33(1): 7-12
- Bombardier C. Outcome assessments in the evaluation of treatment of spinal disorders: summary and general recommendations. *Spine* 2000; 25(24): 3100-3103
- Broadhurst NA & Bond MJ. Pain provocation tests for the assessment of sacroiliac joint dysfunction. *J Spinal Disord*. 1998; 11(4): 341-345
- Burholt V & Nash P. Short Form 36 (SF-36) Health Survey Questionnaire: normative data for Wales. *J. Public Health (Oxf)* 2011; 33(4): 587-603
- Campolina AG, Pinheiro MM, Ciconelli RM & Ferraz MB. Quality of life among the Brazilian adult population using the generic SF-8 questionnaire. *Cad. Saude Publica* 2011; 27(6): 1121-1131
- Carson AJ, Ringbauer B, Stone J et al. Do medically unexplained symptoms matter? A prospective cohort study of 300 new referrals to neurology outpatient clinics. *J. Neurol. Neurosurg. Psychiatry* 2000; 68(2): 207-210
- Casarett D, Karlawish J, Sankar P, Hirschman K & Asch DA. Designing pain research from the patient's perspective: what trial end points are important to patients with chronic pain? *Pain Med* 2001; 2(4): 309-316
- Damen L, Buyruk HM, Guler-Uysal F et al. Pelvic pain during pregnancy is associated with asymmetric laxity of the sacroiliac joints. *Acta Obstet. Gynecol. Scand*. 2001; 80(11): 1019-1024
- Devet, HCW, Terwee, CB, Mokkink, LB & Knol, DL *Measurement in Medicine*. New York: Cambridge University Press; 2011.
- Deyo RA, Battie M, Beurskens AJ et al. Outcome measures for low back pain research. A proposal for standardized use. *Spine* 1998; 23(18): 2003-2013
- Eggen MH, Stuge B, Mowinckel P, Jensen KS & Hagen KB. Can supervised group exercises including ergonomic advice reduce the prevalence and severity of low back pain and pelvic girdle pain in pregnancy? A randomized controlled trial. *Phys. Ther.* 2012; 92(6): 781-790

- Endresen EH. Pelvic pain and low back pain in pregnant women--an epidemiological study. *Scand.J Rheumatol.* 1995; 24(3): 135-141
- Engel GL. The need for a new medical model: a challenge for biomedicine. *Science* 1977; 196(4286): 129-136
- Fagevik OM, Gutke A, Elden H et al. Self-administered tests as a screening procedure for pregnancy-related pelvic girdle pain. *Eur.Spine J.* 2009;
- Fairbank JC & Pynsent PB. The Oswestry Disability Index. *Spine* 2000; 25(22): 2940-2953
- Fairbank JCT, Davies JB, Couper J & O'Brien JP. The Oswestry Low Back Pain Disability Questionnaire. *Physiotherapy* 1980; 66(8): 271-273
- Grotle M, Brox JI & Vollestad NK. Cross-cultural adaptation of the norwegian versions of the roland morris disability questionnaire and the oswestry disability index. *J Rehabil Med* 2003; 35 241-247
- Grotle M, Brox JI & Vøllestad NK. Concurrent comparison of responsiveness in pain and functional status measurements used for patients with low back pain. *Spine* 2004a; 29(21): E492-E501
- Grotle M, Brox JI & Vøllestad NK. Functional status and disability questionnaires: what do they assess? A systematic review of back-specific outcome questionnaires. *Spine* 2004b; Accepted
- Grotle M, Garratt AM, Krogstad JH & Stuge B. Reliability and construct validity of self-report questionnaires for patients with pelvic girdle pain. *Phys.Ther.* 2012; 92(1): 111-123
- Grue, L. Motstand og mestring, om funksjonshemning og livsvilkår. Oslo: Abstract Forlag; 2001.
- Guillemin F., Bombardier C. & Beaton D. Cross-cultural adaptation of health-related quality-of-life measures: literature review and proposed guidelines. *J Clin Epidemiol* 1993; 46 1417-1432
- Gutke A, Hansson ER, Zetherstrom G & Ostgaard HC. Posterior pelvic pain provocation test is negative in patients with lumbar herniated discs. *Eur.Spine J.* 2009;
- Gutke A, Ostgaard HC & Oberg B. Pelvic girdle pain and lumbar pain in pregnancy: A cohort study of the consequences in terms of health and functioning. *Spine* 2006; 31(5): E149-E155
- Gutke A, Ostgaard HC & Oberg B. Predicting persistent pregnancy-related low back pain. *Spine* 2008; 33(12): E386-E393
- Haugso A & Larsen BI. [ICF--a new international classification of functions and health]. *Tidsskr.Nor Laegeforen.* 2004; 124(19): 2514

- Hewlett S, Smith AP & Kirwan JR. Values for function in rheumatoid arthritis: patients, professionals, and public. *Ann.Rheum.Dis.* 2001; 60(10): 928-933
- Kankaanpää M, Taimela S, Laaksonen D, Hanninen O & Airaksinen O. Back and hip extensor fatigability in chronic low back pain patients and controls. *Arch.Phys.Med.Rehabil.* 1998; 79(4): 412-417
- Kankaanpää M, Taimela S, Webber CL, Jr., Airaksinen O & Hanninen O. Lumbar paraspinal muscle fatigability in repetitive isoinertial loading: EMG spectral indices, Borg scale and endurance time. *Eur.J.Appl.Physiol Occup.Physiol* 1997; 76(3): 236-242
- Kerr JM, Vujnovich AL & Bradnam L. Changes in alpha-motoneuron excitability with positions that tension neural tissue. *Electromyogr.Clin.Neurophysiol.* 2002; 42(8): 459-471
- Kjeken I, Kvien TK & Dagfinrud H. [Functional assessment in rehabilitation]. *Tidsskr.Nor Laegeforen.* 2007; 127(5): 598-599
- Kogstad O & Biørnstad N. [Pelvic girdle relaxation. Pathogenesis, etiology, definition, epidemiology]. *Tidsskr.Nor Laegeforen.* 1990; 110(17): 2209-2211
- Kopec JA, Esdaile JM, Abrahamowicz M et al. The Quebec Back Pain Disability Scale. Measurement properties. *Spine* 1995; 20(3): 341-352
- Kristiansson P & Svardsudd K. Discriminatory power of tests applied in back pain during pregnancy. *Spine* 1996; 21(20): 2337-2343
- Kvien TK & Heiberg T. Patient perspective in outcome assessments--perceptions or something more? *J.Rheumatol.* 2003; 30(4): 873-876
- Laslett M, Aprill CN, McDonald B & Young SB. Diagnosis of sacroiliac joint pain: validity of individual provocation tests and composites of tests. *Man.Ther.* 2005a; 10(3): 207-218
- Laslett M, Aprill CN, McDonald B & Young SB. Diagnosis of sacroiliac joint pain: validity of individual provocation tests and composites of tests. *Manual Therapy* 2005b; in press
- Laslett M, Young SB, Aprill CN & McDonald B. Diagnosing painful sacroiliac joints: A validity study of a McKenzie evaluation and sacroiliac provocation tests. *Australian Journal of Physiotherapy* 2003; 49(2): 89-97
- Lefante JJ, Jr., Harmon GN, Ashby KM, Barnard D & Webber LS. Use of the SF-8 to assess health-related quality of life for a chronically ill, low-income population participating in the Central Louisiana Medication Access Program (CMAP). *Qual.Life Res.* 2005; 14(3): 665-673
- Loge JH & Kaasa S. Short form 36 (SF-36) health survey: normative data from the general Norwegian population. *Scand.J.Soc.Med.* 1998; 26(4): 250-258

- Lurie J. A review of generic health status measures in patients with low back pain. *Spine* 2000; 25(24): 3125-3129
- Mens JM, Huis In 't Veld YH & Pool-Goudzwaard A. The Active Straight Leg Raise test in lumbopelvic pain during pregnancy. *Man.Ther.* 2012; 17(4): 364-368
- Mens JM, Pool-Goudzwaard A & Stam HJ. Mobility of the pelvic joints in pregnancy-related lumbopelvic pain: a systematic review. *Obstet.Gynecol.Surv.* 2009; 64(3): 200-208
- Mens JM, Snijders CJ & Stam HJ. Diagonal trunk muscle exercises in peripartum pelvic pain: a randomized clinical trial. *Phys.Ther.* 2000; 80(12): 1164-1173
- Mens JM, Vleeming A, Snijders CJ, Koes BW & Stam HJ. Reliability and validity of the active straight leg raise test in posterior pelvic pain since pregnancy. *Spine* 2001; 26(10): 1167-1171
- Mens JM, Vleeming A, Snijders CJ et al. Responsiveness of outcome measurements in rehabilitation of patients with posterior pelvic pain since pregnancy. *Spine* 2002; 27(10): 1110-1115
- Mens JM, Vleeming A, Snijders CJ, Stam HJ & Ginai AZ. The active straight leg raising test and mobility of the pelvic joints. *Eur.Spine J.* 1999; 8(6): 468-474
- Mens JM, Vleeming A, Stoeckart R, Stam HJ & Snijders CJ. Understanding peripartum pelvic pain. Implications of a patient survey. *Spine* 1996; 21(11): 1363-1369
- Moen MH, Kogstad O, Biørnstad N, Hansen JH & Sudmann E. [Symptomatic pelvic girdle relaxation. Clinical aspects]. *Tidsskr.Nor Laegeforen.* 1990; 110(17): 2211-2212
- Mokkink LB, Terwee CB, Patrick DL et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J.Clin.Epidemiol.* 2010; 63(7): 737-745
- Nilsson-Wikmar L, Holm K, Oijerstedt R & Harms-Ringdahl K. Effect of three different physical therapy treatments on pain and activity in pregnant women with pelvic girdle pain: A randomized clinical trial with 3, 6, and 12 months follow-up postpartum. *Spine* 2005; 30(8): 850-856
- Noren L, Ostgaard S, Nielsen TF & Ostgaard HC. Reduction of sick leave for lumbar back and posterior pelvic pain in pregnancy. *Spine* 1997; 22(18): 2157-2160
- Olsson C & Nilsson-Wikmar L. Health-related quality of life and physical ability among pregnant women with and without back pain in late pregnancy. *Acta Obstet.Gynecol.Scand.* 2004; 83(4): 351-357

- Ostgaard HC, Roos-Hansson E & Zetherstrom G. Regression of back and posterior pelvic pain after pregnancy. *Spine* 1996; 21(23): 2777-2780
- Ostgaard HC, Zetherstrom G & Roos-Hansson E. The posterior pelvic pain provocation test in pregnant women. *Eur.Spine J* 1994; 3(5): 258-260
- Padua L, Padua R, Bondi R et al. Patient-oriented assessment of back pain in pregnancy. *European Spine Journal* 2002; 11(3): 272-275
- Palsson TS & Graven-Nielsen T. Experimental pelvic pain facilitates pain provocation tests and causes regional hyperalgesia. *Pain* 2012;
- Portney, LG & Watkins, MP. *Foundations of clinical research: applications to practice*. Upper Saddle River, N.J: Prentice Hall; 2000.
- Resnik L & Dobrykowski E. Outcomes measurement for patients with low back pain. *Orthop.Nurs.* 2005; 24(1): 14-24
- Robinson HS, Mengshoel AM, Bjelland EK & Vollestad NK. Pelvic girdle pain, clinical tests and disability in late pregnancy. *Man.Ther.* 2010a; 15(3): 280-285
- Robinson HS, Mengshoel AM, Veierod MB & Vollestad N. Pelvic girdle pain: Potential risk factors in pregnancy in relation to disability and pain intensity three months postpartum. *Man.Ther.* 2010b; 15(6): 522-528
- Roe C, Sveen U, Kristoffersen OJ et al. [Testing of ICF core set for low back pain]. *Tidsskr.Nor Laegeforen.* 2008; 128(23): 2706-2708
- Roland M & Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. *Spine* 2000; 25(24): 3115-3124
- Salèn BA, Spangfort EV, Nygren ÅL & Nordemar R. The Disability Rating Index: An Instrument for the Assessment of Disability in Clinical Settings. *J Clin Epidemiol* 1994; 47(12): 1423-1434
- Sosial og helsedirektoratet. Internasjonal klassifisering av funksjon, funksjonshemming og helse. Trondheim: 2003.
- Steiner WA, Ryser L, Huber E et al. Use of the ICF model as a clinical problem-solving tool in physical therapy and rehabilitation medicine. *Phys.Ther.* 2002; 82(11): 1098-1107
- Storheim K, Brox JI, Lochting I, Werner EL & Grotle M. Cross-cultural adaptation and validation of the Norwegian version of the Core Outcome Measures Index for low back pain. *Eur.Spine J.* 2012; 21(12): 2539-2549
- Strender LE, Sjoblom A, Sundell K, Ludwig R & Taube A. Interexaminer reliability in physical examination of patients with low back pain. *Spine* 1997; 22(7): 814-820

Stuge B. [Diagnosis and treatment of pelvic girdle pain]. Tidsskr.Nor Laegeforen. 2010; 130(21): 2141-2145

Stuge B, Garratt A, Krogstad JH & Grotle M. The Pelvic Girdle Questionnaire: A Condition-Specific Instrument for Assessing Activity Limitations and Symptoms in People With Pelvic Girdle Pain. Physical Therapy 2011;

Stuge B, Lærum E, Kirkesola G & Vøllestad N. The efficacy of a treatment program focusing on specific stabilizing exercises for pelvic girdle pain after pregnancy. A randomized controlled trial. Spine 2004a; 29(10): 351-359

Stuge B, Sætre K & Braekken IH. The association between pelvic floor muscle function and pelvic girdle pain--a matched case control 3D ultrasound study. Man.Ther. 2012; 17(2): 150-156

Stuge B, Veierød MB, Lærum E & Vøllestad N. The efficacy of a treatment program focusing on specific stabilizing exercises for pelvic girdle pain after pregnancy. A Two-Year Follow-up of a Randomized Clinical Trial. Spine 2004b; 29(10): E197-E203

Sturesson B, Uden G & Uden A. Pain pattern in pregnancy and "catching" of the leg in pregnant women with posterior pelvic pain. Spine 1997; 22(16): 1880-1883

Terwee CB, Bot SD, de Boer MR et al. Quality criteria were proposed for measurement properties of health status questionnaires. J.Clin.Epidemiol. 2007; 60(1): 34-42

Tonosu J, Takeshita K, Hara N et al. The normative score and the cut-off value of the Oswestry Disability Index (ODI). Eur.Spine J. 2012; 21(8): 1596-1602

Torstensson T, Lindgren A & Kristiansson P. Corticosteroid injection treatment to the ischiadic spine reduced pain in women with long-lasting sacral low back pain with onset during pregnancy: a randomized, double blind, controlled trial. Spine (Phila Pa 1976.) 2009; 34(21): 2254-2258

VAGLUM, P. (2007) Pasienten din har ikke bare en sykdom. Om den biopsykososiale modellen. IN A. Finset & P. Vaglum (Eds.) *Helse, sykdom og atferd. innføring i medisinske atferdsfag*, Oslo, Cappelen Akademisk Forlag.

Vleeming A, Albert HB, Ostgaard HC, Sturesson B & Stuge B. European guidelines for the diagnosis and treatment of pelvic girdle pain. European Spine Journal 2008; 17(6): 794-819

Vleeming A, Vries HJ, Mens JM & van Wingerden JP. Possible role of the long dorsal sacroiliac ligament in women with peripartum pelvic pain. Acta Obstet.Gynecol.Scand. 2002; 81(5): 430-436

Vollestad NK & Stuge B. Prognostic factors for recovery from postpartum pelvic girdle pain. Eur.Spine J. 2009; 18(5): 718-726

Vollestad NK, Torjesen PA & Robinson HS. Association between the serum levels of relaxin and responses to the active straight leg raise test in pregnancy. *Man.Ther.* 2012; 17(3): 225-230

Wade DT & Halligan P. New wine in old bottles: the WHO ICF as an explanatory model of human behaviour. *Clin.Rehabil.* 2003; 17(4): 349-354

Wade DT & Halligan PW. Do biomedical models of illness make for good healthcare systems? *BMJ* 2004; 329(7479): 1398-1401

Ware, JE, Kosinski, M, Dewey, M & et al How to Score and Interpret Single-Item Health Status Measures: a Manual for Users of the SF-8 Health Surcey. Lincoln,R1:QualityMetric Inc; 2001.

Ware JE, Jr. & Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med.Care* 1992; 30(6): 473-483

Wedenberg K, Moen B & Norling A. A prospective randomized study comparing acupuncture with physiotherapy for low-back and pelvic pain in pregnancy. *Acta Obstet.Gynecol.Scand.* 2000; 79(5): 331-335

Wu WH, Meijer OG, Uegaki K et al. Pregnancy-related pelvic girdle pain (PPP), I: Terminology, clinical presentation, and prevalence. *Eur Spine J* 2004; 13 575-589

Vedlegg

Vedlegg 1. Pelvic Girdle Questionnaire (PGQ)

Vedlegg 2. Oswestry liste for funksjonsbegrensninger, versjon 2.0

Vedlegg 3. SF- 8 Health Survey

PGQ

I hvilken grad finner du det problematisk på grunn av plager fra bekkenet, å utføre aktivitetene som er listet opp nedenfor.

Sett et kryss for hver aktivitet som best beskriver hvordan du har det nå for tiden.

Hvor problematisk er det på grunn av bekkenet å:	Ikke i det hele tatt. 0	I liten grad 1	I noen grad 2	I stor grad 3
1.Kle på deg selv				
2.Stå mindre enn 10 minutter				
3.Stå mer enn 60 minutter				
4.Bøye deg				
5.Sitte mindre enn 10 minutter				
6.Sitte mer enn 60 minutter				
7.Gå mindre enn 10 minutter				
8.Gå mer enn 60 minutter				
9.Gå trapper				
10.Husarbeid				
11.Bære lett				
12.Løfte tungt				
13.Reise/sette seg				
14.Skyve en vogn				
15.Løpe				
16.Utføre sportslige aktiviteter*				<input type="checkbox"/>
17.Ligge				
18.Snu deg i sengen				
19.Ha et normalt seksualliv*				<input type="checkbox"/>
20. Skyve noe med den ene foten				

* Hvis ikke aktuelt, kryss av i boksen til høyre

Hvor sterke smerter har du om-	Ingen 0	Noe 1	Moderate 2	Svært mye 3
Morgen				
Kvelden				

I hvilken grad på grunn av plagene i bekkenet,	Ikke i det hele tatt 0	I liten grad 1	I noen grad 2	I stor grad 3
- svikter benet/bena under deg?				
- gjør du ting langsommere?				
- forstyrres nattesøvnen din?				

Scoring procedure: The scores are summarized and recalculated to a percentage score from 0 (no problem at all) to 100 (maximum problem).

OSWESTRY LISTE FOR FUNKSJONSBEGRENSNINGER, versjon 2.0

Vennligst les: Dette spørreskjemaet er utformet for å gi behandleren opplysninger om hvordan ryggsmertene dine har påvirket din evne til å klare deg i dagliglivet. Vennligst svar på hvert avsnitt, og marker bare **det ene feltet** i hvert avsnitt som gjelder for deg. Vi forstår at du kanskje synes at to av utsagnene i hvert avsnitt gjelder deg, men vennligst **marker bare feltet som best beskriver ditt nåværende problem**.

Del 1 – Smerteintensitet

1. Jeg har ingen smerter for øyeblikket
2. Smertene er veldig svake for øyeblikket
3. Smertene er moderate for øyeblikket
4. Smertene er temmelig sterke for øyeblikket
5. Smertene er veldig sterke for øyeblikket
6. Smertene er de verste jeg kan tenke meg for øyeblikket

Del 2 – Personlig stell (vaske seg, kle på seg, osv.)

1. Jeg kan stelle meg selv på vanlig måte uten at det forårsaker ekstra smerter
2. Jeg kan stelle meg selv på vanlig måte, men det er veldig smertefullt
3. Det er smertefullt å stelle meg selv, og jeg gjør det langsomt og forsiktig
4. Jeg trenger noe hjelp, men klarer det meste av mitt personlige stell
5. Jeg trenger hjelp hver dag til det meste av eget stell
6. Jeg kler ikke på meg, har vanskeligheter med å vaske meg, og holder sengen

Del 3 – Løfte

1. Jeg kan løfte tunge ting uten å få mer smerter
2. Jeg kan løfte tunge ting, men får mer smerter
3. Smertene hindrer meg i å løfte tunge ting opp fra gulvet, men jeg greier det hvis det som skal løftes er gunstig plassert, f.eks. på et bord
4. Smertene hindrer meg i å løfte tunge ting, men jeg kan klare lette eller middels tunge ting, hvis det er gunstig plassert
5. Jeg kan bare løfte noe som er veldig lett
6. Jeg kan ikke løfte eller bære noe i det hele tatt

Del 4 – Gå

1. Smerter hindrer meg ikke i å gå i det hele tatt
2. Smerter hindrer meg i å gå mer enn 1 ½ km
3. Smerter hindrer meg i å gå mer enn ¾ km
4. Smerter hindrer meg i å gå mer enn 100 m
5. Jeg kan bare gå med stikk eller krykker
6. Jeg ligger for det meste i sengen og jeg må krabbe til toalettet

Del 5 – Sitte

1. Jeg kan sitte så lenge jeg vil i en hvilken som helst stol
2. Jeg kan sitte så lenge jeg vil i min favorittstol
3. Smerter hindrer meg i å sitte i mer enn en time
4. Smerter hindrer meg i å sitte i mer enn en halv time
5. Smerter hindrer meg i å sitte i mer enn ti minutter
6. Smerter hindrer meg i å sitte i det hele tatt

Forts.OSWESTRY LISTE FOR FUNKSJONSBEGRENSNINGER

Del 6 – Stå

1. Jeg kan stå så lenge jeg vil uten å få mer smerter
2. Jeg kan stå så lenge jeg vil, men får mer smerter
3. Smerter hindrer meg i å stå i mer enn en time
4. Smerter hindrer meg i å stå i mer enn en halv time
5. Smerter hindrer meg i å stå i mer enn ti minutter
6. Smerter hindrer meg i å stå i det hele tatt

Del 7 – Sove

1. Søvn min forstyrres aldri av smerter
2. Søvn min forstyrres av og til av smerter
3. På grunn av smerter får jeg mindre enn seks timers søvn
4. På grunn av smerter får jeg mindre enn fire timers søvn
5. På grunn av smerter får jeg mindre enn to timers søvn
6. Smerter hindrer all søvn

Del 8 – Seksualliv

1. Seksuallivet mitt er normalt og forårsaker ikke mer smerter
2. Seksuallivet mitt er normalt, men forårsaker noe mer smerter
3. Seksuallivet mitt er normalt, men svært smertefullt
4. Seksuallivet mitt er svært begrenset av smerter
5. Seksuallivet mitt er nesten borte på grunn av smerter
6. Smerter forhindrer alt seksualliv

Del 9 – Sosialt liv

1. Det sosiale livet mitt er normalt og forårsaker ikke mer smerter
2. Det sosiale livet mitt er normalt, men øker graden av smerter
3. Smerter har ingen betydelig innvirkning på mitt sosiale liv, bortsett fra at de begrenser mine mer fysiske aktive sider, som sport osv.
4. Smerter har begrenset mitt sosiale liv og jeg går ikke så ofte ut
5. Smerter har begrenset mitt sosiale liv til hjemmet
6. På grunn av smerter har jeg ikke noe sosialt liv

Del 10 – Reising

1. Jeg kan reise hvor som helst uten smerter
2. Jeg kan reise hvor som helst, men det gir mer smerter
3. Smertene er ille, men jeg klarer reiser på to timer
4. Smerter begrenser meg til korte reiser på under en time
5. Smerter begrenser meg til korte, nødvendige reiser på under 30 minutter
6. Smerter forhindrer meg fra å reise, unntatt for å få behandling

Skåring; kode om spørsmålene til 0-5 (1=0...6=5). Summer hvert spørsmål, deles med antall besvarte spørsmål, multipliseres med 0,2 og 100 for å få en prosentskår.

The Modified Oswestry Disability Index (Baker et al 1990)
Oversatt av Margreth Grotle og Nina K.Vøllestad 2001,
Seksjon for Helsefag, Universitetet i Oslo



Din Helse og Trivsel

INTRODUKSJON: Dette spørreskjemaet handler om hvordan du ser på din egen helse. Disse opplysningene vil hjelpe oss til å få vite hvordan du har det og hvordan du er i stand til å utføre dine daglige gjøremål. *Takk for at du fyller ut dette spørreskjemaet.*

For hvert av de følgende spørsmålene vennligst sett et (X) i den ene luken som best beskriver ditt svar.

1. Hvordan vil du, stort sett, vurdere din helsetilstand i løpet av de siste 4 ukene?

☐ Utmerket ☐ Meget god ☐ God ☐ Nokså god ☐ Dårlig ☐ Svært dårlig

2. I løpet av de siste 4 ukene, i hvilken grad begrenset fysiske helseproblemer dine vanlige fysiske aktiviteter (spasere, gå opp trapper)?

☐ Ikke i det hele tatt ☐ Svært lite ☐ En del ☐ Mye ☐ Kunne ikke utføre fysiske aktiviteter

3. I løpet av de siste 4 ukene, hvor vanskelig var det for deg å utføre dittvanlige arbeid (både i og utenfor hjemmet) på grunn av din fysiske helse?

☐ Ikke i det hele tatt ☐ Litt ☐ Nokså ☐ Meget ☐ Kunne ikke utføre daglig arbeid

4. Hvor sterke kroppslige smerter har du hatt i løpet av de siste 4 ukene?

☐ Ingen ☐ Meget svake ☐ Svake ☐ Moderate ☐ Sterke ☐ Meget sterke

5. I løpet av de siste 4 ukene, hvor mye overskudd hadde du?

☐ Svært mye ☐ Ganske mye ☐ En del ☐ Litt ☐ Ikke noe



6. I løpet av de siste 4 ukene, i hvilken grad begrenset din fysiske helse eller følelsesmessige problemer din vanlige sosiale omgang med familie eller venner?

☐ Ikke i det hele tatt ☐ Svært lite ☐ En del ☐ Mye ☐ Kunne ikke ha sosial omgang

7. I løpet av de siste 4 ukene, i hvilken grad har du vært plaget av følelsesmessige problemer (som f.eks. å være engstelig, deprimert eller irritabel)?

☐ Ikke i det hele tatt ☐ Litt ☐ En del ☐ Mye ☐ Svært mye

8. I løpet av de siste 4 ukene, i hvilken grad hindret personlige eller følelsesmessige problemer deg fra å utføre ditt vanlige arbeid, skolegang eller andre daglige gjøremål?

☐ Ikke i det hele tatt ☐ Svært lite ☐ En del ☐ Mye ☐ Kunne ikke utføre daglige gjøremål

Takk for at du svarte på disse spørsmålene!

